PRENTICE-HALL MATHEMATICS SERIES

ALBERT A. BENNETT, EDITOR

Elementary Theory of Equations

BY

William Vernon Lovitt, Ph. D.

PROFESSOR OF MATHEMATICS
COLORADO COLLEGE

PRENTICE-HALL, INC.

ENGLEWOOD CLIFFS, N.J.

First Printing....................November, 1939
Second Printing....................February, 1942
Third Printing........................April, 1946
Fourth Printing....................January, 1947
Fifth Printing........................April, 1948
Sixth Printing......................October, 1948
Seventh Printing....................March, 1950
Eighth Printing......................March, 1953
Ninth Printing........................June, 1955
Tenth Printing........................June, 1957
Eleventh Printing..................January, 1959
Twelfth Printing......................May, 1960
Thirteenth Printing..................April, 1962

PRINTED IN THE UNITED STATES OF AMERICA

26152-C

PREFACE

Heretofore the theory of equations has frequently been presented in a form suitable mainly for graduate and advanced undergraduate students. The present text makes the subject available to those who have completed a one-semester course in Analytic Geometry, and who have not had Calculus.

More material has been included than would ordinarily be given to any one class. The abundance of material provides a choice of subject matter to meet the needs of different teachers who may have different ideas as to what they would like to include in such a course. Also, with a choice of material an instructor may vary from year to year the assignments to be given.

Both in arrangement and wording the author has taken care to make the reading of the text as easy as possible. He firmly believes that many things which appear difficult can be made easy with proper wording. With this fact in mind, explanations are given in full and are not abbreviated.

An unusually large amount of problem material will be found, most of which is new, having been made up and solved in classes taught by the author. Graphical solutions are given for the quadratic and cubic. For the quartic, a graphical solution is presented in the case of two real roots and two imaginary roots. For the first time in a text of this kind Graeffe's method of solution is given.

The author desires here to acknowledge his indebtedness to Prof. A. A. Bennett of Brown University for his kindly criticism and helpful suggestions, which have materially aided in the preparation of this volume.

W. V. LOVITT

CONTENTS

CONTENTS

Elementary Theory of Equations

Chapter I

INTRODUCTION

1.1 General remarks. Every year the number of applications of mathematics to the sciences increases. In these applications equations demanding solution are constantly presented. The equations may be of the simple algebraic type; or they may contain trigonometric or other nonalgebraic functions. They may involve the derivative of an unknown function and be called differential equations. Other possibilities also exist.

Mathematicians have devised methods for solving these equations. Some of these methods lead to exact solutions; others give approximations. Mechanical means have been devised for solving some equations.

The material in this text will deal with algebraic equations and with sets of linear equations. We have made one exception: we introduce a few simple transcendental equations to show the generality and power of Newton's method of solution. Some limit to the material included must be set. Simultaneous algebraic equations of degree higher than the first will not be considered.

Algebraic equations of degree three and four (the cubic and quartic) can be solved by algebraic means. Solutions are given for these two equations. The general algebraic equation of higher degree cannot be solved exactly by radicals. The proof of this statement is, however, beyond the scope of this text.

It is helpful in finding the numerical value of a real root by methods of approximation first to find an interval within which that root lies and in which there is no other root. Hence some simple methods are given for isolating the real roots of an equation.

1.2 Linear equation in one unknown. The equation

$$ax + b = 0, \qquad (a \neq 0)$$

has for solution

$$x = -\frac{b}{a}.$$

Algebraically this is of little interest, although arithmetically it does involve the question of highest common factor. For example, the equation

$$533x - 697 = 0$$

should have its solution presented in simplified form. We have

$$x = \frac{697}{533} = \frac{17}{13}.$$

To make this reduction, we desire to know the *highest common factor** of 697 and 533. To obtain the H.C.F. of two numbers there is a definite method of procedure whereby, barring numerical errors, one cannot fail to find the desired H.C.F. For these two numbers the process is illustrated without comment. A general explanation is given in the next article. 41 is the H.C.F.

533|697|1
533

492
41|164|4
164

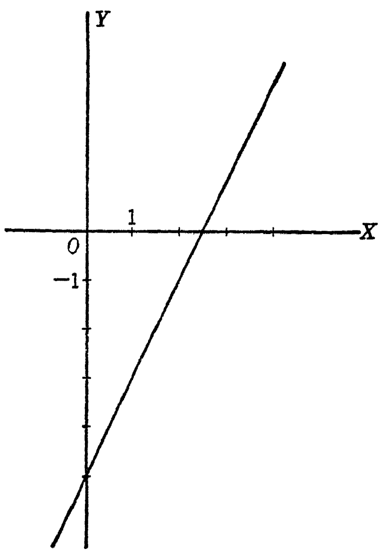


Fig. 1

The real root of $ax + b = 0$ (a and b real, $a \neq 0$) is the x coordinate of the point where the straight line whose equation is $y = ax + b$ cuts the x-axis. For example, the graph of $y = 2x - 5$ cuts the x-axis at $x = 5/2$, as shown in the accompanying graph (Fig. 1).

1.3 Highest common factor. Let A and B ($B \not\equiv 0$) represent two polynomials arranged in descending powers of some common unknown letter (such as x). Let the degree of A be equal to or

* Usually written H.C.F.

greater than that of B. If Q represents the quotient and R_1 the remainder on dividing A by B, then

$$A \equiv BQ + R_1; \tag{1}$$

$$\therefore R_1 \equiv A - BQ. \tag{2}$$

From (1), we see that A is exactly divisible by every factor common to B and R_1. From (2), we see that R_1 is exactly divisible by every factor common to A and B. Hence, the common factors of A and B are the common factors of B and R_1. Divide B by R_1. Let the quotient be Q_1 and the remainder R_2. At each step continue the division until the remainder is of a degree lower than the divisor. Continuing in this way, we have

$$\left.\begin{aligned} A &\equiv BQ + R_1; \\ B &\equiv R_1Q_1 + R_2; \end{aligned}\right\} \tag{3}$$

$$R_{i-1} \equiv R_iQ_i + c \qquad (c, \text{ constant}).$$

If $c \neq 0$, A and B have no common factor containing x. For it follows from the identities (3) that A and B have the same common factors as B and R_1; B and R_1 have the same common factors as R_1 and R_2, and finally R_{i-1} and R_i have the same common factors as R_i and c. But since c is a constant (not zero), R_i and c have no common factor containing the unknown. Hence A and B can have no common factor containing the unknown.

If $c = 0$, A and B have a common factor R_i containing x explicitly. For every factor of R_i is a factor of R_{i-1}, and R_i is the common factor of highest degree. But the common factors of R_i and R_{i-1} are common factors of R_{i-1} and R_{i-2}. Hence, R_i is the H.C.F. of R_{i-1} and R_{i-2}. Continuing in this way, we see that R_i is the H.C.F. of A and B. When A or B contains a literal coefficient, the equation $c = 0$ becomes the condition under which A and B have a common factor containing x.

In order to avoid fractions, we may, in any of the divisions, multiply or divide the dividend or divisor by a constant. This will change any subsequent remainder at most by a constant multiplier and hence the H.C.F. at most by a constant multiplier.

Example 1. Find the H.C.F. of $x^3 + 1$ and $x^4 + x^3 - x^2 + 2x$.

$$
\begin{array}{l|llll}
\underline{B = x^3 + 1} & A = x^4 + x^3 - x^2 + 2x & \underline{|\,x + 1} & & \\
 & \underline{x^4 \qquad\qquad + x} & & & \\
 & \quad x^3 - x^2 + x & & & \\
 & \quad \underline{x^3 \qquad\quad + 1} & & & \\
R_1 = & \quad \underline{-x^2 + x - 1} & |\,x^3 \qquad + 1 & \underline{|-x - 1} & \\
 & & \underline{x^3 - x^2 + x} & & \\
 & & \quad x^2 - x + 1 & & \\
 & & \quad \underline{x^2 - x + 1} & & \\
 & & \quad R_2 = c = 0 & &
\end{array}
$$

The H.C.F. is $x^2 - x + 1$.

Exercises

Solve these linear equations, giving the answer in its simplest form:

1. $671x - 793 = 0$	5. $1577x - 996 = 0$
2. $632x - 1185 = 0$	6. $1027x - 1896 = 0$
3. $1909x - 1079 = 0$	7. $3186x - 1298 = 0$
4. $1608x - 3350 = 0$	8. $1968x - 1763 = 0$

Find the H.C.F. of the following:

9. $x^3 + x^2 - 2$ and $x^3 + 2x^2 - 4x + 1$
10. $x^4 + x^3 + 2x^2 + x + 1$ and $x^3 - x^2 + x - 1$
11. $x^3 - 3x + 2$ and $x^2 - 1$
12. $x^4 + 5x^3 + 9x^2 + 7x + 2$ and $4x^3 + 15x^2 + 18x + 7$
13. $x^5 + 2x^4 - 2x^3 - 4x^2 + x + 2$ and $x^3 - 7x - 6$
14. $x^5 - x^4 + 8x^2 - 8x$ and $x^3 + 4x^2 - 8x + 24$
15. $x^5 + 2x^4 - 2x^3 - 4x^2 + x + 2$ and $5x^4 + 8x^3 - 6x^2 - 8x + 1$
16. $x^5 - 2x^4 + x^3 - x^2 + 2x - 1$ and $5x^4 - 8x^3 + 3x^2 - 2x + 2$
17. $x^5 + 2x^4 + x^3 + x^2 + 2x + 1$ and $10x^3 + 12x^2 + 3x + 1$
18. $2x^4 + 9x^3 + 14x + 3$ and $3x^4 + 15x^3 + 5x^2 + 10x + 2$
19. $x^5 + 4x^4 - x^3 - 16x^2 - 11x - 2$ and $x^6 + 6x^5 + 10x^4 - 10x^2 - 6x - 1$

In the following determine the condition on the literal coefficients so that the given pair of polynomials shall have a common factor containing x. When the letter k is involved, find k and the corresponding common factor.

20. $x^2 - 4x + 3$ and $x^2 + kx + 4$
21. $x + k$ and $kx + 1$
22. $3ax^2 + 2bx + c$ and $bx^2 + 2cx + 3d$
23. $x^2 + 2x + k$ and $x^2 + 3x + 2k$

1.4 Quadratics. The simple quadratic equation

$$x^2 + a^2 = 0 \qquad (a \neq 0, \text{ real}) \quad (4)$$

has for roots

$$x_1 = +a\sqrt{-1}, \qquad x_2 = -a\sqrt{-1}.$$

$a\sqrt{-1}$ is called a *pure imaginary number.*

The quadratic equation

$$x^2 + x + 1 = 0$$

has for roots

$$x_1 = -\frac{1}{2} + \frac{\sqrt{-3}}{2}, \qquad x_2 = -\frac{1}{2} - \frac{\sqrt{-3}}{2}.$$

These two numbers are called *complex numbers.* Pure imaginary numbers and complex numbers will be given more consideration in Chapter III.

The general quadratic equation

$$ax^2 + bx + c = 0 \qquad (a \neq 0) \quad (5)$$

can be written as the sum of two squares as follows:

$$a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) = 0. \quad (6)$$

Make the substitution $z = x + \dfrac{b}{2a}$.

Equation (6) becomes

$$az^2 + \left(c - \frac{b^2}{4a}\right) = 0. \quad (7)$$

By this process the term of degree one lower than the term of highest degree has been removed. There is a general process for the removal of the term of degree one lower than the term of highest degree. This will be shown in the chapter on transforma-

tion of equations. The removal of the second degree term is one of the first steps in the solution of the cubic as given in Chapter VII.

Equations (4) and (7) are known as *binomial equations*; they contain two terms only. The solution of binomial equations assumes some importance in the Theory of Equations. Their solution will be taken up in Chapter III. There is a theorem, of Abel, the proof of which is beyond the scope of this book, to the effect that *every equation which is solvable by radicals can be reduced to a chain of binomial equations of prime degree whose roots are rational functions of the roots of the given equation.*

The roots of (5) are

$$x_1 = -\frac{b}{2a} + \frac{\sqrt{b^2 - 4ac}}{2a}, \qquad x_2 = -\frac{b}{2a} - \frac{\sqrt{b^2 - 4ac}}{2a}.$$

Whence:

$$\text{the difference of roots } = x_1 - x_2 = \frac{\sqrt{b^2 - 4ac}}{a};$$

$$\text{the sum of roots } = x_1 + x_2 = -\frac{b}{a};$$

$$\text{the product of roots } = x_1 x_2 = \frac{c}{a}.$$

Hence $\sqrt{b^2 - 4ac}$, the only irrational function of the coefficients which occurs in the roots, is rationally expressible in terms of the roots. Furthermore, we see that the symmetric functions of the roots, $x_1 + x_2$ and $x_1 x_2$, are expressible rationally in terms of the coefficients.

$\Delta \equiv b^2 - 4ac$ is known as the *discriminant* of the equation. We have

$$\Delta = a^2(x_1 - x_2)^2;$$

that is, the discriminant, except for the constant factor a^2, is the product of the squares of the differences of the roots.

The equation (7) has the roots

$$z_1 = \frac{\sqrt{b^2 - 4ac}}{2a}, \qquad z_2 = -\frac{\sqrt{b^2 - 4ac}}{2a}.$$

The discriminant of the transformed equation (7) is equal to the discriminant of the original equation (4); for

$$a^2(z_1 - z_2)^2 = b^2 - 4ac = a^2(x_1 - x_2)^2.$$

1.5 Graphical discussion. The real roots of $ax^2 + bx + c = 0$ (a, b, c real, $a \neq 0$) are the x coordinates of the points where the curve

$$y = ax^2 + bx + c \tag{8}$$

crosses the x-axis. It is easy to verify that from (8) we have

$$y = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right]. \tag{9}$$

From (9), y will have its smallest value ($a > 0$) when $\left(x + \dfrac{b}{2a}\right)^2$ has its smallest value; that is, when $x = -\dfrac{b}{2a}$. For this value of x we find $y = \dfrac{4ac - b^2}{4a}$. This is the point P in Fig. 2. P is said to be a minimum point on the curve. The roots of the quadratic are represented by the points A and B.

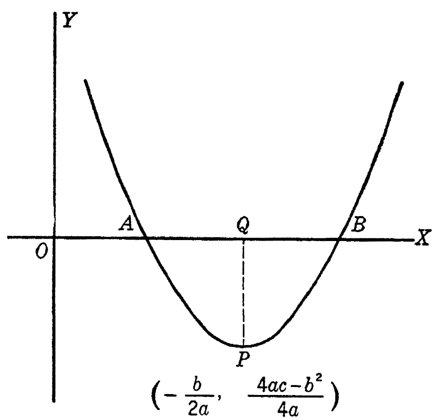


Fig. 2

$$OQ = -\frac{b}{2a}; \qquad AQ = QB = \frac{\sqrt{b^2 - 4ac}}{2a}$$

From (9), y will have its largest value ($a < 0$) when $x = -\dfrac{b}{2a}$. Such a point is said to be a maximum point on the curve. The curve as a whole is shifted vertically, up or down, a distance equal to the change in c; up if c is increased, down if c is decreased. This is illustrated by the three curves in Fig. 3:

(a) $y = x^2 - 2x - 3$, (b) $y = x^2 - 2x + 1$, (c) $y = x^2 - 2x + 2$.

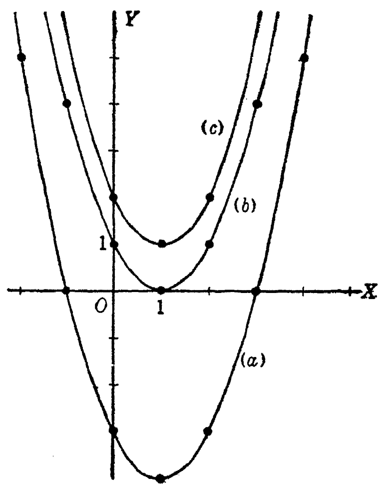


Fig. 3

Curve (a) cuts the x-axis at $x = -1, +3$. The roots of the corresponding quadratic are -1, $+3$, and the discriminant is positive. Curve (b) is tangent to the x-axis at $x = 1$. The corresponding quadratic equation has two equal roots, $+1$, $+1$, and the discriminant is zero. Curve (c) does not intersect the x-axis. The roots of the corresponding quadratic are imaginary, and the discriminant is negative.

The minimum in each case is at $x = -\frac{b}{2a} = 1$. The minimum value of y for (a) is -4; for (b) is 0; for (c) is $+1$.

Exercises

Solve each of the following quadratics by first expressing the left-hand member as the sum of two squares:

1. $x^2 + 6x + 5 = 0$
2. $6x^2 + 13x + 6 = 0$
3. $2x^2 + 3x - 5 = 0$
4. $x^2 + 6x + 9 = 0$
5. $x^2 - x + 1 = 0$
6. $4x^2 - 5x + 2 = 0$

7. For each of the above equations verify the formulas for the sum and product of the roots.

8. For each of the above quadratics verify that

$$x_1 - x_2 = \frac{\sqrt{b^2 - 4ac}}{a}.$$

9. For each of the above quadratics compute the discriminant and verify that $b^2 - 4ac = a^2(x_1 - x_2)^2$.

10. Verify that the discriminant of $ax^2 + bx + c$ vanishes if, and only if, the linear functions $2ax + b$ and $bx + 2c$ are proportional.

Chapter II

SIMULTANEOUS LINEAR EQUATIONS IN TWO AND THREE VARIABLES

2.1 Introduction. We have solved the linear and quadratic equations. The solution of the cubic and quartic will be taken up in a later chapter. The general algebraic equation of degree higher than the fourth cannot be solved by radicals. Approximation methods will be developed later for obtaining the real roots of algebraic equations of any degree.

In this chapter we will consider the solution of simultaneous linear equations in two and three variables. In a later chapter on determinants we will take up the solution of n simultaneous linear equations in m unknowns.

2.2 Two linear equations in two unknowns.

Given the two equations

$$\begin{aligned} a_1x + b_1y &= c_1, \\ a_2x + b_2y &= c_2, \end{aligned} \tag{1}$$

solve for x and y. Multiply the first equation by b_2, the second equation by b_1 and subtract.

We have

$$(a_1b_2 - a_2b_1)x = c_1b_2 - c_2b_1,$$

whence

$$x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}, \qquad (a_1b_2 - a_2b_1 \neq 0).$$

Multiply the first equation by a_2, the second by a_1; subtract and solve for y. We find

$$y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}.$$

We observe that the denominators are the same. The numerator in the expression which gives the value of x can be obtained from

the denominator by replacing the coefficients of x by the constant terms; that is, a_1 is replaced by c_1, and a_2 by c_2. The numerator in the expression which gives the value of y can be obtained from the denominator by replacing the coefficients of y by the constant terms.

The values of x and y may be written as follows:

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \tag{2}$$

The expressions in the numerator and denominator of the fractions in (2), giving the values of x and y, are called *determinants of the second order*.

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1$$

a_1, a_2, b_1, b_2 are called *elements* of the determinant. The elements a_1, b_2 constitute the principal diagonal. The elements a_2, b_1 constitute the secondary diagonal. We obtain the algebraic expression $a_1b_2 - a_2b_1$ for which the determinant stands by multiplying together the elements in the principal diagonal and subtracting the product of the elements in the secondary diagonal.

Equations (2) may be used as formulas, providing a quick solution of equations (1).

Example: Solve $5x + 2y = 16$
$3x - 4y =$

$$x = \frac{\begin{vmatrix} 16 & 2 \\ -6 & -4 \end{vmatrix}}{\begin{vmatrix} 5 & 2 \\ 3 & -4 \end{vmatrix}} = \frac{-64 + 12}{-20 - 6} = \frac{-52}{-26} = 2,$$

$$y = \frac{\begin{vmatrix} 5 & 16 \\ 3 & -6 \end{vmatrix}}{\begin{vmatrix} 5 & 2 \\ 3 & -4 \end{vmatrix}} = \frac{-30 - 48}{-26} = \frac{-78}{-26} = 3.$$

The two equations represent two straight lines which intersect in the point whose coordinates are (2, 3).

If $a_1b_2 - a_2b_1 = 0$, then $a_1 = ka_2$ and $b_1 = kb_2$, where k is a constant, and our equations become

$$ka_2x + kb_2y = c_1,$$

$$a_2x + b_2y = c_2,$$

and the two equations represent parallel (perhaps coincident) straight lines.

If $c_1 \neq kc_2$, the two lines are distinct and the equations are said to be *inconsistent*. If $c_1 = kc_2$, the two lines are coincident and the equations are said to be *dependent*. For example,

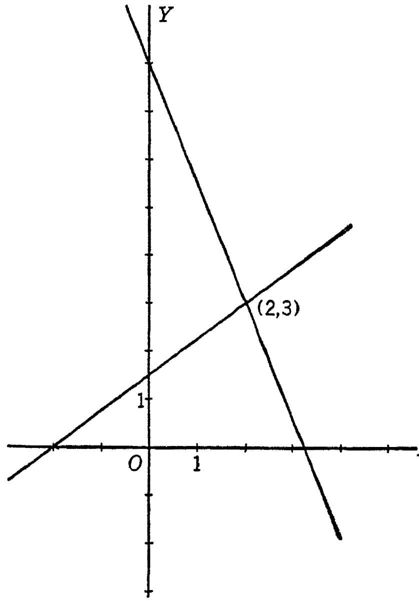


Fig. 4

$$\left.\begin{aligned} 2x + 3y &= 6 \\ 4x + 6y &= 24 \end{aligned}\right\}$$ are inconsistent and represent two distinct parallel lines.

$$\left.\begin{aligned} 2x + 3y &= 6 \\ 4x + 6y &= 12 \end{aligned}\right\}$$ are dependent and represent two coincident straight lines.

Exercises

Evaluate:

1. $\begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix}$

2. $\begin{vmatrix} 2 & -2 \\ 3 & 5 \end{vmatrix}$

3. $\begin{vmatrix} 4 & 3 \\ -2 & 5 \end{vmatrix}$

4. $\begin{vmatrix} 5 & 0 \\ 16 & 2 \end{vmatrix}$

5. $\begin{vmatrix} 3 & 7 \\ 0 & 0 \end{vmatrix}$

6. $\begin{vmatrix} 6 & 3 \\ 2 & 1 \end{vmatrix}$

7. $\begin{vmatrix} a & b \\ 2a & 2b \end{vmatrix}$

8. $\begin{vmatrix} x & y \\ ax & ay \end{vmatrix}$

9. $\begin{vmatrix} 1 & \sin x \\ \sin x & 1 \end{vmatrix}$

10. $\begin{vmatrix} \sec x & \tan x \\ \tan x & \sec x \end{vmatrix}$

11. $\begin{vmatrix} \sin x & -\cos x \\ \cos x & \sin x \end{vmatrix}$

12. $\begin{vmatrix} \csc x & \cot x \\ \cot x & \csc x \end{vmatrix}$

Solve by determinants:

13. $2x + 5y = 20, \qquad 3x + 2y = 19$

14. $5x - 4y = 8, \qquad x + 3y = 13$

15. $5x + 2y = 4, \qquad 4x - 3y = 17$

16. $3x - 2y = 19, \qquad 2x - 3y = 21$

17. $\dfrac{1}{x} + \dfrac{1}{y} = 5, \qquad \dfrac{3}{x} - \dfrac{4}{y} = 1$

18. $\dfrac{2}{x} + \dfrac{1}{y} = 9, \qquad \dfrac{6}{x} - \dfrac{2}{y} = 1$

19. $\dfrac{4}{x} + \dfrac{3}{y} = 4, \qquad \dfrac{2}{x} - \dfrac{6}{y} = 7$

20. $2x + \dfrac{3}{y} = 5, \qquad 3x - \dfrac{6}{y} = 4$

For the quadratic $ax^2 + bx + c$, **show that:**

21. The discriminant is the square of $a \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix}$.

22. The discriminant is $-\begin{vmatrix} 2a & b \\ b & 2c \end{vmatrix}$.

Show that:

23. The area of the triangle formed by $(0, 0)$, (x_1, y_1), (x_2, y_2) is half the absolute value of $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$.

24. $\begin{vmatrix} a - b & a + b \\ c - d & c + d \end{vmatrix} = 2 \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

25. $\begin{vmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}^2$.

2.3 Solution of three linear equations in three unknowns. Solve the following equations for x, y, and z:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1, \\ a_2x + b_2y + c_2z &= d_2, \\ a_3x + b_3y + c_3z &= d_3. \end{aligned} \tag{3}$$

Multiply the first equation by c_2, the second by c_1, and subtract. Multiply the first equation by c_3, the third by c_1, and subtract. We have

$$(a_1c_2 - a_2c_1)x + (b_1c_2 - b_2c_1)y = d_1c_2 - d_2c_1,$$

$$(a_1c_3 - a_3c_1)x + (b_1c_3 - b_3c_1)y = d_1c_3 - d_3c_1.$$

If we solve these last two equations for x and y, we obtain

$$x = \frac{d_1b_2c_3 + d_2b_3c_1 + d_3b_1c_2 - d_3b_2c_1 - d_2b_1c_3 - d_1b_3c_2}{a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2},$$

$$y = \frac{a_1d_2c_3 + a_2d_3c_1 + a_3d_1c_2 - a_3d_2c_1 - a_2d_1c_3 - a_1d_3c_2}{a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2}.$$

In like manner we find

$$z = \frac{a_1b_2d_3 + a_2b_3d_1 + a_3b_1d_2 - a_3b_2d_1 - a_2b_1d_3 - a_1b_3d_2}{a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2},$$

provided the denominator is not zero.

We observe that the denominators are the same. The numerators differ from the denominators only in that the coefficients of the variable under consideration are replaced by the constant terms.

Let us represent the denominator by the symbol

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

This symbol is called a *determinant* of the third order. Several schemes are in use to assist one in recalling the algebraic expression represented by the symbol. One is as follows:

$$\begin{array}{ccc|cc} a_1 & b_1 & c_1 & a_1 & b_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ a_3 & b_3 & c_3 & a_3 & b_3 \end{array}$$

Repeat the first two columns. Draw a diagonal line through $a_1\ b_2\ c_3$. This is one term in the expansion of the determinant. Draw two more diagonal lines parallel to this first diagonal line, each line passing through three elements of the determinant. In this way we obtain three terms, all positive, namely: $a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3$.

Now draw a secondary diagonal line through $a_3\ b_2\ c_1$. This is a negative term in the expansion. Two more diagonal lines are drawn parallel to this secondary diagonal, each line passing through three elements of the determinant. In this way we obtain three terms, all negative: $-a_3b_2c_1$; $-b_3c_2a_1$; $-c_3a_2b_1$. No similar diagram exists for a determinant of higher order.

With this notation, the solution of equations (3) can be expressed formally as follows:

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}; \quad y = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}; \quad z = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} \tag{4}$$

Example: Solve

$$\begin{aligned} x + y + z &= 6 \\ 2x + y - 2z &= 6 \\ 3x - y - 3z &= 4 \end{aligned}$$

$$x = \frac{\begin{vmatrix} 6 & 1 & 1 \\ 6 & 1 & -2 \\ 4 & -1 & -3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & -2 \\ 3 & -1 & -3 \end{vmatrix}}; \quad y = \frac{\begin{vmatrix} 1 & 6 & 1 \\ 2 & 6 & -2 \\ 3 & 4 & -3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & -2 \\ 3 & -1 & -3 \end{vmatrix}}; \quad z = \frac{\begin{vmatrix} 1 & 1 & 6 \\ 2 & 1 & 6 \\ 3 & -1 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & -2 \\ 3 & -1 & -3 \end{vmatrix}}$$

But

$$\left|\begin{matrix} 1 & 1 & 1 \\ 2 & 1 & -2 \\ 3 & -1 & -3 \end{matrix}\right|\begin{matrix} 1 & 1 \\ 2 & 1 \\ 3 & -1 \end{matrix}\Bigg| \qquad -3 - 6 - 2 - 3 - 2 + 6 = -10;$$

and

$$\left|\begin{matrix} 6 & 1 & 1 \\ 6 & 1 & -2 \\ 4 & -1 & -3 \end{matrix}\right|\begin{matrix} 6 & 1 \\ 6 & 1 \\ 4 & -1 \end{matrix}\Bigg| = -18 - 8 - 6 - 4 - 12 + 18 = -30,$$

whence

$$x = \frac{-30}{-10} = 3. \quad \text{In like manner } y = 2,\ z = 1.$$

Exercises

Evaluate the following determinants:

1. $\begin{vmatrix} 3 & 5 & 1 \\ 2 & 1 & 4 \\ 1 & 2 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 5 & -1 & 2 \\ 2 & 3 & -4 \\ -1 & 2 & 3 \end{vmatrix}$

3. $\begin{vmatrix} 6 & 9 & 2 \\ 4 & 6 & 2 \\ 2 & 3 & 3 \end{vmatrix}$

4. $\begin{vmatrix} 3 & 2 & 0 \\ 0 & 2 & 5 \\ 1 & 0 & 4 \end{vmatrix}$

5. $\begin{vmatrix} 5 & -4 & 3 \\ 2 & 3 & -6 \\ -1 & 2 & 1 \end{vmatrix}$

6. $\begin{vmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{vmatrix}$

7. $\begin{vmatrix} 10 & 20 & 30 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$

Solve the equations below by determinants:

9. $\begin{aligned} x + y + z &= 10 \\ 2x + 3y - 2z &= 10 \\ 3x - y - 3z &= 4 \end{aligned}$

10. $\begin{aligned} 3x - 2y + z &= 17 \\ 2x + y - 2z &= 4 \\ x + y - 4z &= -2 \end{aligned}$

11. $\begin{aligned} x + y + z &= 1 \\ 2x + 3y - 4z &= -4 \\ 3x + 5y + 4z &= 3 \end{aligned}$

12. $\begin{aligned} x + y - z &= 0 \\ 2x - y + z &= 3 \\ 5x + 2y - 3z &= 0 \end{aligned}$

13. $\begin{aligned} 2x + y - 2z &= 1 \\ x + y - z &= 0 \\ 3x - 4z &= 1 \end{aligned}$

14. $\begin{aligned} x + y &= 2 \\ y + z &= 1 \\ x - z &= 1 \end{aligned}$

15. $\begin{aligned} x + y - 2z &= 0 \\ 2x + 3y + 2z &= 5 \\ 3y - 4z &= 4 \end{aligned}$

16. $\begin{aligned} x + y + 2z &= 0 \\ 2x - 3y + 2z &= 2 \\ x + z &= 4 \end{aligned}$

Show that:

17. $\begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix} = -(x_2 - x_3)(x_3 - x_1)(x_1 - x_2).$

18. The quadratic equation whose roots are x_1 and x_2 $(x_1 \neq x_2)$ may be written as

$$\begin{vmatrix} x^2 & x & 1 \\ x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \end{vmatrix} = 0$$

19. $$\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix} = 0$$

20. $$\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix} = (a+b+c)(a^2+b^2+c^2-bc-ac-ab)$$

21. The equation of the straight line through (x_1, y_1) and (x_2, y_2) is given by $\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$

2.4 Rank of D. Notation. Let

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix};\ N_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix};\ N_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix};\ N_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

Any determinant of the second order, formed from the elements of D by deleting a horizontal row and a vertical column of elements is called a *two-rowed minor* of D; likewise for N_x, N_y, N_z. D is said to be of rank 1 if every two-rowed minor is equal to zero, but some element is not equal to zero.

2.5 Geometric interpretation. All of the coefficients in (3) will be considered real.

Case I. $D \neq 0$. The equations (3) represent three planes which intersect in a real point.

Case II. $D = 0$ *and is of rank* 1, *but some element not zero*; N_x, N_y, N_z *of rank* 1.

The three planes coincide. For from

$$\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = \begin{vmatrix} a_i & c_i \\ a_j & c_j \end{vmatrix} = \begin{vmatrix} b_i & c_i \\ b_j & c_j \end{vmatrix} = \begin{vmatrix} a_i & d_i \\ a_j & d_j \end{vmatrix} = \begin{vmatrix} b_i & d_i \\ b_j & d_j \end{vmatrix}$$
$$= \begin{vmatrix} c_i & d_i \\ c_j & d_j \end{vmatrix} = 0 \quad (i, j = 1, 2, 3) \tag{5}$$

it follows that

$$a_i = k_i a_1, \quad b_i = k_i b_1, \quad c_i = k_i c_1, \quad d_i = k_i d_1 \quad (i = 2, 3), \tag{6}$$

and the three planes coincide.

Case III. $D = 0$ *and of rank* 1, *but some element not zero, and*

$$\begin{vmatrix} a_i & d_i \\ a_j & d_j \end{vmatrix}, \quad \begin{vmatrix} b_i & d_i \\ b_j & d_j \end{vmatrix}, \quad \begin{vmatrix} c_i & d_i \\ c_j & d_j \end{vmatrix} \text{ not all zero} \quad (i, j = 1, 2, 3) \tag{7}$$

The three planes are parallel and not all coincident.
For from

$$\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = \begin{vmatrix} a_i & c_i \\ a_j & c_j \end{vmatrix} = \begin{vmatrix} b_i & c_i \\ b_j & c_j \end{vmatrix} = 0 \quad (i, j = 1, 2, 3) \tag{8}$$

it follows that

$$a_i = k_i a_1, \quad b_i = k_i b_1, \quad c_i = k_i c_1 \quad (i = 2, 3) \tag{9}$$

and hence the three planes are parallel. From (7) and (9) not both of the following can be true:

$$d_2 = k_2 d_1, \quad d_3 = k_3 d_1;$$

and hence not all three planes coincide.

Case IV. $D = 0$, *but some two-rowed minor* $\neq 0$, *and* $N_x = N_y = N_z = 0$.

The three planes have a common line of intersection.

Let $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$, for example, be the two-rowed minor of D that is not zero. There exist constants k_1 and k_2 such that

$$k_1 b_1 + k_2 b_2 = b_3 \quad \text{and} \quad k_1 c_1 + k_2 c_2 = c_3. \tag{10}$$

To k_1 times the elements in the first row add k_2 times the corresponding elements in the second row and subtract the sums so obtained from the corresponding elements in the third row. We have

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \overset{*}{=} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 - k_1 a_1 - k_2 a_2 & 0 & 0 \end{vmatrix}$$

$$= (a_3 - k_1 a_1 - k_2 a_2) \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} = 0.$$

* The justification for this transformation is given in a later chapter on determinants (§11.16).

By hypothesis:

$$\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \neq 0,$$

therefore

$$a_3 = k_1a_1 + k_2a_2. \tag{11}$$

Applying the same transformation to $N_x = 0$, we find

$$d_3 = k_1d_1 + k_2d_2. \tag{12}$$

From (10), (11), (12), it follows that $N_y = N_z = 0$.
Now (10), (11), (12) assure us that the third equation is a linear combination of the first two. Therefore the three planes either intersect in a common real line or are parallel. If they were parallel D would need to be of rank 1 as indicated in Case III. But D is not of rank 1, since by hypothesis some two-rowed minor $\neq 0$. Therefore the planes intersect in a common real line.

Case V. $D = 0$ but some two-rowed minor $\neq 0$. N_x, N_y, N_z not all zero.

Two of the planes intersect in a line parallel to the third plane. Equations (4) can be written as follows:

$$Dx = N_x, \qquad Dy = N_y, \qquad Dz = N_z.$$

But $D = 0$, and N_x, N_y, N_z are not all zero; therefore equations (3) are inconsistent.

A logical process of elimination of possibilities assures us that the given conditions leave us the only remaining configuration of three planes, namely that stated above. This configuration can happen in two ways:

(1) Two parallel planes are cut by a third plane.

(2) The three planes intersect in three mutually parallel lines.

Exercises

Discuss the following sets of equations:

1. $\begin{aligned} x + y + z &= 3 \\ 2x - 2y + 2z &= 2 \\ 3x - y + 3z &= 2 \end{aligned}$

2. $\begin{aligned} x + y + 2z &= 7 \\ 2x - 2y + z &= 3 \\ 3x - y + 2z &= 9 \end{aligned}$

3. $$\begin{aligned} x + y + z &= 3 \\ 2x + 2y + 2z &= 2 \\ 3x - y - 3z &= 5 \end{aligned}$$

4. $$\begin{aligned} 3x - 2y + z &= -6 \\ 2x + 5y - 3z &= 2 \\ 4x - 9y + 5z &= -14 \end{aligned}$$

5. $$\begin{aligned} 2x + y - z &= 3 \\ 4x + 2y - 2z &= 6 \\ 6x + 3y - 3z &= 9 \end{aligned}$$

6. $$\begin{aligned} x + 2y - z &= 1 \\ 2x + 4y - 2z &= 1 \\ 3x + 6y - 3z &= 2 \end{aligned}$$

7. $$\begin{aligned} x + 2y + 3z &= 6 \\ 2x + 3y + 4z &= 8 \\ 3x + 4y + 5z &= 10 \end{aligned}$$

8. $$\begin{aligned} 6x + 9y + 2z &= 1 \\ 4x + 6y + 2z &= 2 \\ 2x + 3y + 3z &= 3 \end{aligned}$$

9. $$\begin{aligned} 2x + y + 5z &= 4 \\ x - 3y + z &= 5 \\ x + 4y + 4z &= -1 \end{aligned}$$

10. $$\begin{aligned} 3x + 2y &= -6 \\ x - y - 3z &= -2 \\ 6x - y - 9z &= 0 \end{aligned}$$

11. $$\begin{aligned} 2x - y + 3z &= 1 \\ 3x + 2y - 2z &= 2 \\ 7y - 13z &= 1 \end{aligned}$$

12. $$\begin{aligned} 5x + 2y - 3Z &= 4 \\ 2x - 3y + 2Z &= 5 \\ 4x + 13y - 12Z &= -7 \end{aligned}$$

13. $$\begin{aligned} 6x + 9y - 12Z &= \\ 8x + 12y - 16Z &= 9 \\ 4x + 6y - 8Z &= 7 \end{aligned}$$

14. $$\begin{aligned} 2x + 3y - 4Z &= 1 \\ 4x + 6y - 8Z &= 3 \\ x + y + Z &= 2 \end{aligned}$$

15. $$\begin{aligned} 4x - 3y &= -1 \\ 2y - 4Z &= 2 \\ 2x - y - Z &= 3 \end{aligned}$$

16. $$\begin{aligned} 4x + 6y + Z &= 4 \\ 2x + 3y - 2Z &= 2 \\ 6x + 9y + 3Z &= 3 \end{aligned}$$

17. $$\begin{aligned} 3x - 2y + 4Z &= 1 \\ 9x - 6y + Z &= 3 \\ 6x - 4y - 2Z &= 4 \end{aligned}$$

18. $$\begin{aligned} x + y - Z &= 2 \\ 2x - 3y + 2Z &= 3 \\ x + 11y - 9Z &= 4 \end{aligned}$$

19. $$\begin{aligned} 5x - y + 2Z &= 4 \\ 10x - 2y + 4Z &= 3 \\ x + y + Z &= 1 \end{aligned}$$

20. $$\begin{aligned} 3x + y - 2Z &= 1 \\ 3x + y - 2Z &= 2 \\ 6x + 2y - 4Z &= 4 \end{aligned}$$

21. $$\begin{aligned} 4x + 2y + 6Z &= 8 \\ 6x + 3y + 9Z &= 12 \\ 8x + 4y + 12Z &= 16 \end{aligned}$$

22. $$\begin{aligned} x + y + Z &= 4 \\ 2x - 3y - 4Z &= 1 \\ x - 2y + 2Z &= 9 \end{aligned}$$

Chapter III

SOLUTION OF BINOMIAL EQUATIONS

3.1 Introduction. We saw in Chapter I, in solving the quadratic, that complex numbers occur. We also discovered that a quadratic equation could be reduced to a binomial equation. It was stated without proof that *every equation which is solvable by radicals can be reduced to a chain of binomial equations of prime degree whose roots are rational functions of the roots of the given equation.*

The general binomial equation is of the form:

$ax^{n+r} + bx^r = 0$ ($a \neq 0, b \neq 0$; r an integer $r \geq 0$) or $x^r(ax^n + b) = 0$.

The factor x^r gives rise to the root 0 with multiplicity r. It remains to solve

$$ax^n + b = 0, \qquad (a \neq 0, b \neq 0). \tag{1}$$

The trivial case $b = 0$ gives a root 0 with multiplicity n. Since by hypothesis $a \neq 0$, there is no loss of generality in confining our attention to the case $a = 1$, since other cases may be reduced to this by dividing by a. Hereafter in referring to a binomial equation we shall mean an equation of the form (1) with n a positive integer, $a = 1$.

This chapter will be devoted to setting up the machinery for the solution of a binomial equation and effecting the solution.

3.2 Pure Imaginary Numbers. Formally the solution of $x^2 + a^2 = 0$ is given by $x = \pm a\sqrt{-1}$.

It is customary to let the letter i represent the *imaginary unit* $\sqrt{-1}$. This imaginary unit has the property that its square is -1. We have

$$\begin{aligned}
i^2 &= -1, & i^5 &= i \cdot i^4 = i(+1) = +i,\\
i^3 &= i \cdot i^2 = i(-1) = -i, & i^6 &= i \cdot i^5 = i \cdot i \quad = -1,\\
i^4 &= i \cdot i^3 = i(-i) = +1, & i^7 &= i \cdot i^6 = i(-1) = -i, \text{ etc.}
\end{aligned}$$

The solution of $x^2 = -a^2$ is usually written $x = \pm ai$.

The solution of $x^2 = -9$ is $x = \pm 3i$.

The solution of $x^2 = -c$ is usually written $x = \pm\sqrt{c}\,i$, and not $x = \pm\sqrt{-c}$, although either is correct.

A pure imaginary number is the product of a real number and this imaginary unit. Thus, $2i$, $3i$ and even 0, are pure imaginary numbers. The square of a pure imaginary number is always a negative real number or zero.

Thus,

$$(2i)^2 = -4;\ (-3i)^2 = (-3)^2(i)^2 = 9i^2 = -9.$$

3.3 Complex Numbers. A *complex number* is the sum of a real number and a pure imaginary number. Thus, $3 + 2i$, $4 - 3i$, $-3 + 5i$, $a + bi$ (for a and b real) and even 2 and i are complex numbers. Any complex number of the form $a + bi$, where $b \neq 0$, and which, therefore, is not merely a real number (namely a), is called an *imaginary* number. The pure imaginary numbers (with $a = 0$) clearly are but special cases of these imaginary numbers. $a - bi$ is said to be the *conjugate* of $a + bi$. $a + bi = 0$, if, and only if, $a = b = 0$. Otherwise we would have $a = -bi$, that is, a pure imaginary number equal to a real number, which is impossible unless both are zero. Two complex numbers $a + bi$ and $c + di$ are equal if, and only if, $a = c$ and $b = d$. Otherwise we would have $a - c = i(d - b)$, which cannot be, for $a - c$ is real and $i(d - b)$ is imaginary, unless $a - c$ and $b - d$ are both zero.

Operations with complex numbers are performed exactly as in the field of formal algebra of real numbers, save only that $i^2 = -1$. Thus,

$$\begin{array}{lll}
\text{I} & (a + bi) + (c + di) & = (a + c) + (b + d)i \\
 & (4 + 3i) + (3 + 2i) & = 7 + 5i \\[6pt]
\text{II} & (a + bi) - (c + di) & = (a - c) + (b - d)i \\
 & (4 + 3i) - (3 + 2i) & = 1 + i \\[6pt]
\text{III} & (a + bi)(c + di) & = (ac - bd) + (bc + ad)i \\
 & (4 + 3i)(3 + 2i) & = 6 + 17i
\end{array}$$

$$\text{IV}\quad \frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\,i$$

$$\frac{4 + 3i}{3 + 2i} = \frac{(4 + 3i)(3 - 2i)}{(3 + 2i)(3 - 2i)} = \frac{18 + i}{13} = \frac{18}{13} + \frac{1}{13}\,i$$

Exercises

Express as complex numbers:

1. $(5 + 6i) + (3 + 2i)$
2. $(5 + 6i) + (3 - 4i)$
3. $(4 + 3i) + (4 - 3i)$
4. $(4 + 3i) - (4 - 3i)$
5. $(4 + 3i)(3 + 4i)$
6. $(3 + 2i)(2 + 3i)$
7. $(5 + 7i)8i$
8. $(x + iy)^2$
9. $(5 + 3i)^2$
10. $(5 - 3i)^2$
11. $(3 + 2i)^3$
12. $(2 + 3i)^3$
13. $\dfrac{3 + 5i}{2 + i}$
14. $\dfrac{6 + 5i}{2 - 3i}$
15. $\dfrac{4 + 3i}{3 - 4i}$
16. $\dfrac{2 + 5i}{2 + 3i}$
17. $\dfrac{7 + 3i}{2 - 5i}$
18. $\dfrac{5 + 3i}{4 + i}$

19. Find real numbers x and y for which

$$(x + iy)^2 = 5 + 12i$$

Solution: $x^2 - y^2 + 2xyi = 5 + 12i$;

therefore $x^2 - y^2 = 5$ and $xy = 6$,

whence $x = 3, \quad y = 2$ or $x = -3, \quad y = -2$.

20. Find real numbers x and y for which

(a) $(x + iy)^2 = 21 + 20i$, (b) $(x + iy)^2 = 40 + 42i$,
(c) $(x + iy)^2 = -5 + 12i$, (d) $(x + iy)^2 = -16 + 30i$.

21. Express as complex numbers the square roots of

(a) $5 - 12i$, (b) $7 + 24i$, (c) $9 + 40i$.

22. Prove that the conjugate of the sum of two complex numbers is equal to the sum of their conjugates.

23. Prove that the conjugate of the product of two complex numbers is equal to the product of their conjugates.

24. Prove that the conjugate of the quotient of two complex numbers is equal to the quotient of their conjugates.

25. Prove that a complex number is equal to its conjugate if, and only if, it is real.

26. Prove that the sum of a complex number and its conjugate is zero if, and only if, the number is a pure imaginary.

27. Prove that the product of conjugate complex numbers is real, and not negative.

3.4 Geometrical representation of complex numbers.

Any complex number $a + bi$ can be represented by a point P whose coordinates are (a, b), and the point P in turn identifies the complex number. This is called the *rectangular form* of the complex number. Thus there is a one-to-one correspondence between the points of a plane and complex numbers. Real numbers are represented by the points on the x-axis, which is called the *axis of reals*. Pure imaginary numbers are represented by the points on the y-axis, which is called the *axis of imaginaries*. Real numbers and pure imaginary numbers are thus exhibited as special cases of complex numbers.

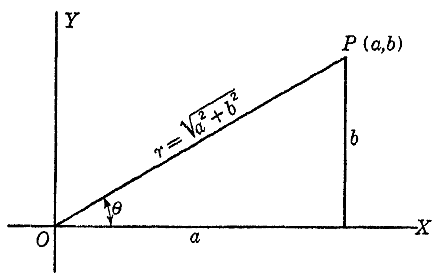


Fig. 5

The directed line OP, instead of the point P, may be considered the geometric representation of the complex number. A directed line segment such as OP is called a *vector*, here drawn from the origin O. It represents both a length and a direction. From figure 5 it is evident that

$$a = r \cos \theta, \quad b = r \sin \theta, \quad r = \sqrt{a^2 + b^2}.$$

Then

$$a + bi = r(\cos \theta + i \sin \theta).$$

This is the *trigonometric form* of the complex number. The angle θ is called the *amplitude* (or *argument*) of the complex number; r is called the *modulus** (or *absolute value*) of $a + bi$. $\cos \theta + i \sin \theta$ is sometimes called a *complex unit* and is sometimes indicated by the notation *cis* θ.

The magnitude and direction of a *force* can be represented by the length and direction of a line segment. It follows that a complex number $a + bi$, or its equivalent $r(\cos \theta + i \sin \theta)$, can

* Plural, *moduli*.

represent a force whose magnitude is $r = \sqrt{a^2 + b^2}$ and whose direction is $\theta = \arctan \frac{b}{a}$.

Exercises

1. Plot the following complex numbers:

a. $2 + 3i$	e. $6 + 8i$	j. $\frac{1}{2} + \frac{\sqrt{3}}{2} i$
b. $2 - 3i$	f. $12 - 5i$	k. $\frac{1}{2} - \frac{\sqrt{3}}{2} i$
c. $4 + 3i$	g. $5i$	l. $-\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i$
d. $3 - 4i$	h. $-5i$	
	i. $-3 + 2i$	

2. Find the amplitude and modulus of each of the numbers in exercise 1.

3. Plot the following complex numbers:

a. $\cos 45° + i \sin 45°$	g. $\cos 225° + i \sin 225°$
b. $\cos 30° + i \sin 30°$	h. $\cos 300° + i \sin 300°$
c. $\cos 120° + i \sin 120°$	i. $2(\cos 30° + i \sin 30°)$
d. $\cos 150° + i \sin 150°$	j. $3(\cos 135° + i \sin 135°)$
e. $\cos 210° + i \sin 210°$	k. $4(\cos 330° + i \sin 330°)$
f. $\cos 240° + i \sin 240°$	l. $2(\cos 390° + i \sin 390°)$

4. Solve the following equations and plot the roots:

a. $x^2 + x + 1 = 0$	f. $x^2 + 4x + 13 = 0$
b. $x^2 - x + 1 = 0$	g. $x^3 - 1 = 0$
c. $x^2 - 6x + 13 = 0$	h. $x^4 - 1 = 0$
d. $x^2 + 6x + 13 = 0$	i. $x^3 + 1 = 0$
e. $x^2 - 4x + 13 = 0$	j. $x^6 - 1 = 0$

3.5 Graphical solution of the quadratic.

In Chapter I we saw that $ax^2 + bx + c$ $(a \neq 0)$ can be put in the form $a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right]$. Whence the lowest (highest) point on the curve $y = ax^2 + bx + c$ is for $x = -b/2a$. The corresponding value of y is $(4ac - b^2)/4a$. The roots of $ax^2 + bx + c = 0$ are

$$x_1 = -\frac{b}{2a} + \frac{\sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad x_2 = -\frac{b}{2a} - \frac{\sqrt{b^2 - 4ac}}{2a}$$

If $b^2 - 4ac < 0$, both roots are imaginary, say $\alpha \pm \beta i$. Then $\alpha = -b/2a = OA$ (Fig. 6), and $\beta^2 = (4ac - b^2)/4a^2 = \overline{AM}/a$. Since a and AM are known, β can now be determined graphically.

The lines $y = m\left(x + \dfrac{b}{2a}\right)$ through A (Fig. 6) with slopes $\pm\sqrt{4ac - b^2}$ are tangents to the curve $y = ax^2 + bx + c$. If the points of tangency are P and R, and Q is the mid-point of RP, then $RQ = QP = \dfrac{\sqrt{4ac - b^2}}{2a} = \beta$. Mark off $AB = B'A = RQ = \beta$. Then the points B and B', thought of as being in the complex plane, represent the roots $\alpha + \beta i$ and $\alpha - \beta i$ respectively.

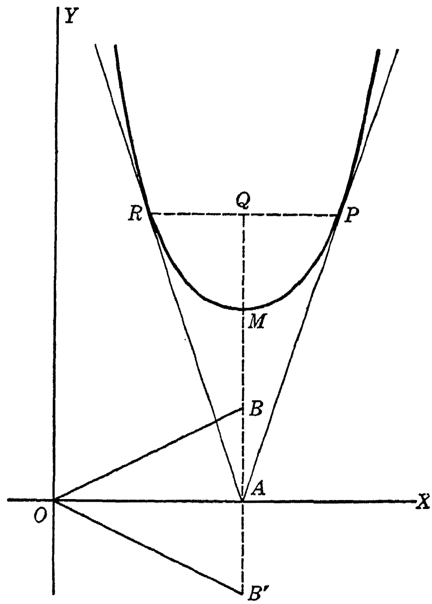


Fig. 6

3.6 Multiplication of complex numbers. Let z_1 and z_2 be any two complex numbers:

$$z_1 = r_1 (\cos\theta_1 + i \sin\theta_1);$$
$$z_2 = r_2 (\cos\theta_2 + i \sin\theta_2).$$

Then

$$\begin{aligned} z_1z_2 &= r_1 (\cos\theta_1 + i\sin\theta_1)\cdot r_2 (\cos\theta_2 + i\sin\theta_2) \\ &= r_1r_2 [(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)] \\ &= r_1r_2 [\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]. \end{aligned}$$

We now observe that *the product of two complex numbers is another complex number whose modulus is the product of the moduli of the two factors and whose amplitude is the sum of the amplitudes of the factors.*

If there is a third factor $z_3 = r_3 (\cos\theta_3 + i\sin\theta_3)$, then

$$\begin{aligned} z_1z_2z_3 &= r_1r_2\,[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)]\cdot r_3(\cos\theta_3 + i\sin\theta_3) \\ &= r_1r_2r_3\,[\cos(\theta_1+\theta_2+\theta_3) + i\sin(\theta_1+\theta_2+\theta_3)]. \end{aligned} \tag{1}$$

If there are n factors, then

$$\begin{aligned} z_1z_2\cdots z_n &= r_1(\cos\theta_1 + i\sin\theta_1)\cdot r_2(\cos\theta_2 + i\sin\theta_2)\cdots \\ &\qquad r_n(\cos\theta_n + i\sin\theta_n) \\ &= r_1r_2\cdots r_n\,[\cos(\theta_1+\theta_2+\cdots+\theta_n) \\ &\qquad + i\sin(\theta_1+\theta_2+\cdots+\theta_n)]. \end{aligned} \tag{2}$$

De Moivre's theorem. In (2) put $r_1 = r_2 = \cdots = r_n = r$ and $\theta_1 = \theta_2 = \cdots = \theta_n = \theta$. We have

$$[r(\cos\theta + i\sin\theta)]^n = r^n(\cos n\theta + i\sin n\theta).$$

In particular, if $r = 1$, we have

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta. \tag{3}$$

For n not thus restricted to positive integers this is known as De Moivre's theorem. It was discovered in 1730.

We have proved De Moivre's theorem for n, a positive integer. We shall now complete the proof of this theorem for any rational value of n.*

(a) Let n be a fraction $\frac{r}{s}$, r and s being positive integers.

By (3) we have

$$\cos\theta + i\sin\theta = \left(\cos\frac{\theta}{s} + i\sin\frac{\theta}{s}\right)^s.$$

Take the s-th root of both sides. We have

$$(\cos\theta + i\sin\theta)^{\frac{1}{s}} = \cos\frac{\theta}{s} + i\sin\frac{\theta}{s}.$$

Raising both sides to the r-th power, we have

$$(\cos\theta + i\sin\theta)^{\frac{r}{s}} = \cos\frac{r}{s}\theta + i\sin\frac{r}{s}\theta.$$

(b) Let n be any negative rational number, $-m$, for example: -2, or $-5/7$. Then

* Proofs for the cases in which n is not rational lie beyond the scope of this volume.

$$\begin{aligned}(\cos\theta + i\sin\theta)^n &= (\cos\theta + i\sin\theta)^{-m}\\ &= \frac{1}{(\cos\theta + i\sin\theta)^m} = \frac{1}{\cos m\theta + i\sin m\theta}\\ &= \frac{\cos m\theta - i\sin m\theta}{\cos^2 m\theta + \sin^2 m\theta} = \cos m\theta - i\sin m\theta\\ &= \cos(-m\theta) + i\sin(-m\theta)\\ &= \cos n\theta + i\sin n\theta.\end{aligned}$$

3.7 Powers of a complex number. De Moivre's theorem enables us to raise a complex number to any integral power, or to extract any root of a complex number. The complex number must, however, be put in trigonometric form. The procedure is illustrated by the examples which follow:

Example 1. Find the cube of $1 + i$.

$$1 + i = \sqrt{2}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = \sqrt{2}\,(\cos 45^\circ + i\sin 45^\circ).$$

Then

$$\begin{aligned}(1+i)^3 &= [\sqrt{2}\,(\cos 45^\circ + i\sin 45^\circ)]^3\\ &= 2\sqrt{2}\,(\cos 135^\circ + i\sin 135^\circ)\\ &= 2\sqrt{2}\left(-\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = -2 + 2i.\end{aligned}$$

Example 2. Raise $4 + 3i$ to the fourth power.

$$\begin{aligned}4 + 3i &= 5\left(\frac{4}{5} + i\frac{3}{5}\right)\\ &= 5(\cos 36^\circ\ 52'\ 12'' + i\sin 36^\circ\ 52'\ 12'').\end{aligned}$$

Then

$$\begin{aligned}(4+3i)^4 &= [5\,(\cos 36^\circ\,52'\,12'' + i\sin 36^\circ\,52'\,12'')]^4\\ &= 625\,(\cos 147^\circ\,28'\,48'' + i\sin 147^\circ\,28'\,48'')\\ &= 625\,(-0.84321 + i\,0.53759)\\ &= -527 + 336i.\end{aligned}$$

Example 3. Cube $a + bi$

$$\begin{aligned}a + bi &= \sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}} + i\frac{b}{\sqrt{a^2+b^2}}\right)\\ &= \sqrt{a^2+b^2}\,(\cos\theta + i\sin\theta)\end{aligned}$$

where $$\theta = \arctan \frac{b}{a}.$$

Then
$$(a + bi)^3 = [\sqrt{a^2 + b^2}(\cos\theta + i\sin\theta)]^3$$
$$= (a^2 + b^2)^{3/2}(\cos 3\theta + i \sin 3\theta).$$

Now determine the numerical value of $\sin 3\theta$ and $\cos 3\theta$ and simplify.

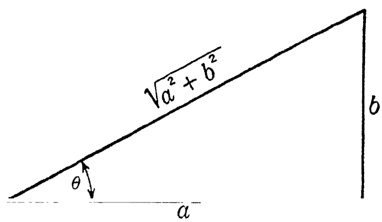


Fig. 7

Exercises

1. Express the following numbers in the form $r(\cos\theta + i\sin\theta)$:

a. 3
b. -2
c. $3i$
d. $-4i$
e. $1 + i$
f. $1 - i$
g. $1 + \sqrt{3}\,i$
h. $1 - \sqrt{3}\,i$
i. $-1 + \sqrt{3}\,i$
j. $\sqrt{3} - i$
k. $\sqrt{3} + i$
l. $-1 - \sqrt{3}\,i$

Find by De Moivre's theorem the value of

2. $(\cos 10° + i\sin 10°)^3$
3. $(\cos 15° + i\sin 15°)^4$
4. $(\cos 20° + i\sin 20°)^9$
5. $(\cos 120° + i\sin 120°)^{-2}$
6. $(\cos 30° + i\sin 30°)^{12}$
7. $(\cos 45° + i\sin 45°)^{17}$
8. $(\cos 120° + i\sin 120°)^3$
9. $(\cos 240° + i\sin 240°)^3$
10. $(\cos 60° + i\sin 60°)^{31}$
11. $(\cos 135° + i\sin 135°)^{27}$
12. $(\cos 30° + i\sin 30°)^{-1}$
13. $(\cos 60° + i\sin 60°)^{-3}$
14. $(\cos 135° + i\sin 135°)^{2/3}$
15. $(\cos 60° + i\sin 60°)^{3/2}$

16. a. 3^2
b. $(2i)^4$
c. $(1 + i)^5$
d. $(1 - i)^3$
e. $(\sqrt{3} + i)^4$
f. $(-\sqrt{3} + i)^5$
g. $(3 + \sqrt{3}\,i)^6$
h. $(2 + 2\sqrt{3}\,i)^3$
i. $(-2\sqrt{3} + 2i)^4$
j. $\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)^3$
k. $\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right)^3$
l. $\left(\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}\right)^4$
m. $(-\sqrt{-3})^5$
n. $(3 - 4i)^3$
o. $(12 + 5i)^4$

3.8 Roots of a complex number. In order to extract a root of a complex number, first put it in the trigonometric type form

$$r\,(\cos\theta + i\sin\theta).$$

Then, since

$$\cos\theta = \cos(2k\pi + \theta) \quad \text{and} \quad \sin\theta = \sin(2k\pi + \theta),$$

where k is an integer, we can place the complex number in the equivalent form,

$$r\,[\cos(2k\pi + \theta) + i\sin(2k\pi + \theta)].$$

The method is illustrated by the following examples.

Example 1. Find the cube roots of -1.

$$\begin{aligned}-1 &= \cos 180° + i\sin 180° \\ &= \cos(2k\pi + 180°) + i\sin(2k\pi + 180°).\end{aligned}$$

Then

$$\begin{aligned}(-1)^{1/3} &= [\cos(2k\pi + 180°) + i\sin(2k\pi + 180°)]^{1/3} \\ &= \cos\frac{2k\pi + 180°}{3} + i\sin\frac{2k\pi + 180°}{3}.\end{aligned}$$

If $k = 0$, then $(-1)^{\frac{1}{3}} = \cos 60° + i\sin 60° = \frac{1}{2} + i\frac{\sqrt{3}}{2}$;

if $k = 1$, then $(-1)^{\frac{1}{3}} = \cos 180° + i\sin 180° = -1$;

if $k = 2$, then $(-1)^{\frac{1}{3}} = \cos 300° + i\sin 300° = \frac{1}{2} - i\frac{\sqrt{3}}{2}$.

Assignment of other values to k yields no new results. The values just found are the roots of the binomial equation $x^3 + 1 = 0$. The two imaginary roots $\frac{1}{2} + i\frac{\sqrt{3}}{2}$ and $\frac{1}{2} - i\frac{\sqrt{3}}{2}$ are roots of the quadratic equation $x^2 - x + 1 = 0$.

Example 2. Evaluate $(-1 - i\sqrt{3})^{1/4}$.

$$\begin{aligned}-1 - i\sqrt{3} &= 2\,(\cos 240° + i\sin 240°) \\ &= 2\,[\cos(2k\pi + 240°) + i\sin(2k\pi + 240°)].\end{aligned}$$

Then

$$(-1 - i\sqrt{3})^{1/4} = 2^{1/4}\left[\cos\frac{2k\pi + 240°}{4} + i\sin\frac{2k\pi + 240°}{4}\right].$$

If $k = 0$, we have $1.1892 (\cos 60° + i \sin 60°) = 0.5946 + 1.0299i$;
if $k = 1$, we have $1.1892 (\cos 150° + i \sin 150°) = -1.0299 + 0.5946i$;
if $k = 2$, we have $1.1892 (\cos 240° + i \sin 240°) = -0.5946 - 1.0299i$;
if $k = 3$, we have $1.1892 (\cos 330° + i \sin 330°) = 1.0299 - 0.5946i$.

Assignment of other values to k yields no new results.

Every time that we find the n roots of a complex number $a + bi$, we have found the roots of the corresponding binomial equation

$$x^n = a + bi.$$

3.9 Geometric representation of the solutions of a binomial equation. The geometric representation of the roots of a binomial equation will be made clear by the following illustrations:

Illustration 1. Solve the binomial equation $x^3 - 1 = 0$ and plot the roots.

Factoring

$$x^3 - 1 = (x - 1)(x^2 + x + 1) = 0,$$

whence

$$x - 1 = 0; \qquad \therefore x = 1$$

$$x^2 + x + 1 = 0; \qquad \therefore x = -\tfrac{1}{2} \pm \frac{\sqrt{3}}{2} i.$$

The three roots of the given equation are

$$x_1 = 1, \qquad x_2 = -\tfrac{1}{2} + \frac{\sqrt{3}}{2} i, \qquad x_3 = -\tfrac{1}{2} - \frac{\sqrt{3}}{2} i.$$

The reader can readily verify that $x_3 = x_2^2$. Then if we put $x_2 = \omega$, we have $x_3 = \omega^2$. The roots of this particular binomial equation are often given as $1, \omega, \omega^2$.

The reader should verify the following relations:

$$\omega^3 = 1, \qquad \omega \cdot \omega^2 = 1, \qquad (\omega^2)^2 = \omega, \qquad (\omega^2)^3 = 1,$$

$$\Sigma x_i = 1 + \omega + \omega^2 = 0; \qquad \Sigma x_i x_j = 1 + \omega + \omega^2 = 0,$$

$$x_1 \cdot x_2 \cdot x_3 = 1.$$

The roots $1, \omega, \omega^2$ are plotted in the accompanying figure 8. The roots are represented by the points A, B, C; and the points A, B, C identify the roots.

If we use De Moivre's theorem to obtain the roots, we have

$$x^3 = 1 = \cos 2k\pi + i \sin 2k\pi,$$

whence

$$x = (\cos 2k\pi + i \sin 2k\pi)^{\frac{1}{3}} = \cos \frac{2k\pi}{3} + i \sin \frac{2k\pi}{3}. \tag{4}$$

Giving k the values 0, 1, 2, we have

$$x_1 = 1 = \cos 0° + i \sin 0°;$$

$$x_2 = \omega = \cos 120° + i \sin 120° = -\tfrac{1}{2} + \frac{\sqrt{3}}{2} i;$$

$$x_3 = \omega^2 = \cos 240° + i \sin 240° = -\tfrac{1}{2} - \frac{\sqrt{3}}{2} i.$$

Notice that in (4) the argument $(2k\pi)$ is a positive integral multiple of 360°. To obtain the amplitude for ω, 360° is divided by three. The amplitude for ω^2 is twice that for ω. That is, the points A, B, C which identify the roots are the trisection points of the circumference of the unit circle. Furthermore the roots of $x^3 = a$ would be equally spaced on a circle with radius $|\sqrt[3]{a}|$. Joining the points A, B, C by straight line segments, one has a regular polygon of three sides.

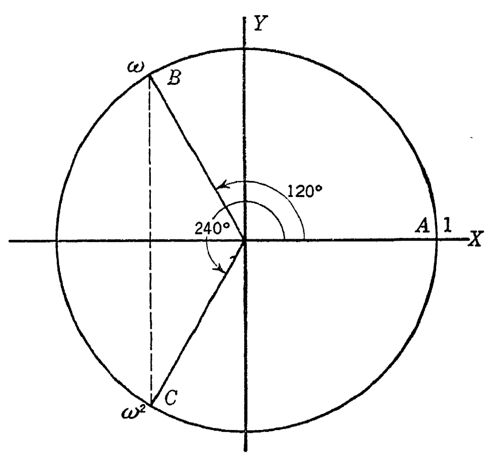


Fig. 8

We have solved by algebraic means the binomial equation $x^3 - 1 = 0$ without the use of any irrationalities other than the

extraction of real square roots. One can construct with straight-edge and compass alone a regular polygon of three sides.

Illustration 2. Solve the binomial equation $x^4 - 1 = 0$ and plot the roots.

Factoring,

$$x^4 - 1 = (x - 1)(x + 1)(x - i)(x + i) = 0.$$

The roots are

$$x_1 = 1; \qquad x_2 = i; \qquad x_3 = -1; \qquad x_4 = -i.$$

Using De Moivre's theorem, we have

$$x^4 = 1 = \cos 2k\pi + i \sin 2k\pi,$$

whence

$$x = (\cos 2k\pi + i \sin 2k\pi)^{1/4} = \cos \frac{2k\pi}{4} + i \sin \frac{2k\pi}{4}. \tag{5}$$

Giving k the values 0, 1, 2, 3, we have

$$\begin{aligned}
x_1 &= \cos 0^\circ + i \sin 0^\circ = 1\\
x_2 &= \cos 90^\circ + i \sin 90^\circ = i\\
x_3 &= \cos 180^\circ + i \sin 180^\circ = -1\\
x_4 &= \cos 270^\circ + i \sin 270^\circ = -i.
\end{aligned}$$

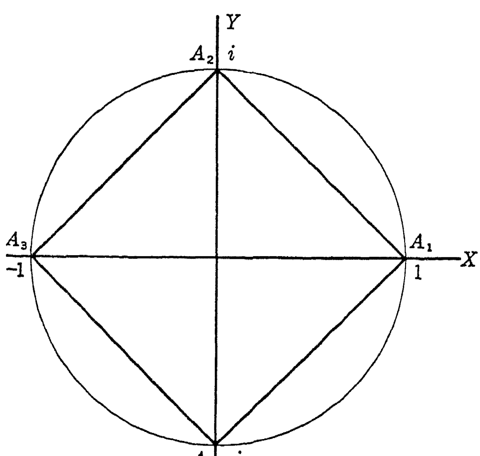


Fig. 9

Notice that in (5) the argument $(2k\pi)$ is a positive integral multiple of 360°. To obtain the amplitude for x_2, 360° is divided by four. The amplitudes for the roots x_3, x_4, x_1 are obtained by multiplying the amplitude for x_2 successively by 2, 3, 4.

Thus we see that the roots $1, i, -1, -i$ represented by the points A_1, A_2, A_3, A_4

in the accompanying figure 9 are equally spaced on the circumference of the unit circle. Furthermore, the roots of $x^4 = a$ would be equally spaced on the circumference of a circle with radius $|\sqrt[4]{a}|$. By joining the points A_1, A_2, A_3, A_4 with straight line segments a regular polygon of four sides results.

We have solved by algebraic means the binomial equation $x^4 - 1 = 0$ without the use of any irrationalities other than the extraction of real square roots. One can construct with straightedge and compass alone a regular polygon of four sides.

Note that

$$x_1^4 = x_2^4 = x_3^4 = x_4^4 = 1;$$

also

$$x_3 = x_2^2, \qquad x_4 = x_2^3, \qquad x_1 = x_2^4.$$

If $x_2 = R$, the roots may be represented by R, R^2, R^3, $R^4 = 1$. The student should verify that $x_2x_4 = 1$; $x_1x_3 = -1$; $\Sigma x_i = 0$; $\Sigma x_ix_j = 0$ $(i \neq j)$; $\Sigma x_ix_jx_k = 0$ $(i \neq j, k; j \neq k)$; $x_1x_2x_3x_4 = -1$.

Illustration 3. Solve the binomial equation $x^6 - 1 = 0$ and plot the roots.

Factoring,

$$x^6 - 1 = (x - 1)(x^2 + x + 1)(x + 1)(x^2 - x + 1) = 0.$$

The roots are

$$x_1 = 1; \qquad x_2 = \tfrac{1}{2} + \frac{\sqrt{3}}{2}i; \qquad x_3 = \omega; \qquad x_4 = -1; \qquad x_5 = \omega^2;$$

$$x_6 = \tfrac{1}{2} - \frac{\sqrt{3}}{2}i.$$

Using De Moivre's theorem, we have

$$x^6 = 1 = \cos 2k\pi + i \sin 2k\pi$$

and

$$x = (\cos 2k\pi + i \sin 2k\pi)^{1/6} = \cos\frac{2k\pi}{6} + i \sin\frac{2k\pi}{6} \tag{6}$$

Giving k the values 0, 1, 2, 3, 4, 5, we have

$$x_1 = \cos 0° + i \sin 0° = 1$$

$$x_2 = \cos 60° + i \sin 60° = \tfrac{1}{2} + \frac{\sqrt{3}}{2}i = -\omega^2$$

$$x_3 = \cos 120° + i \sin 120° = -\tfrac{1}{2} + \frac{\sqrt{3}}{2} i = \omega$$

$$x_4 = \cos 180° + i \sin 120° = -1$$

$$x_5 = \cos 240° + i \sin 240° = -\tfrac{1}{2} - \frac{\sqrt{3}}{2} i = \omega^2$$

$$x_6 = \cos 300° + i \sin 300° = \tfrac{1}{2} - \frac{\sqrt{3}}{2} i = -\omega.$$

Notice that in (6) the argument $(2k\pi)$ is a positive integral multiple of 360°. To obtain x_2, 360° is divided by six. The arguments for the roots x_3, x_4, x_5, x_6, x_1 are obtained by multiplying the argument for x_2 successively by 2, 3, 4, 5, 6. Thus we see that the roots of the binomial equation $x^6 - 1 = 0$ are equally spaced on the circumference of the unit circle. Furthermore, the roots of $x^6 = a$ would be equally spaced on the circumference of a circle with radius $| \sqrt[6]{a} |$.

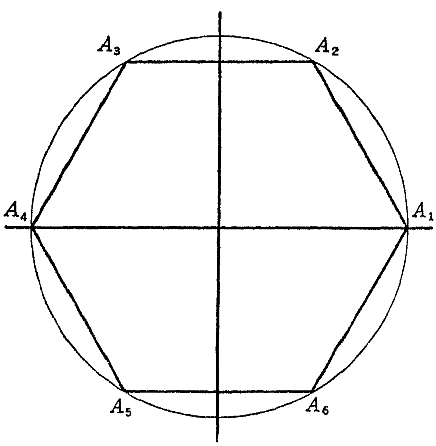


Fig. 10

Let the six roots of $x^6 = 1$ be represented on the unit circle by the points A_i ($i = 1$ to 6); joining these points by straight line segments, one has a regular polygon of six sides.

We have solved by algebraic means the binomial equation $x^6 - 1 = 0$ without the use of any irrationalities other than the extraction of real square roots.

Note that $x_i^6 = 1$ $(i = 1, \cdots, 6)$; also $x_3 = x_2^2$; $x_4 = x_2^3$; $x_5 = x_2^4$; $x_6 = x_2^5$; $x_1 = x_2^6$. If $x_2 = R$, the roots may be represented by $R, R^2, R^3, R^4, R^5, R^6 = 1$.

3.10 Primitive roots of unity. In the general case $x^n - 1 = 0$, we have

$$x^n = 1 = \cos 2k\pi + i \sin 2k\pi$$

and

$$x = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \qquad (k = 0, 1, 2, \cdots, n-1) \quad (7)$$

All of the nth roots of unity are included in (7).

Let R represent the nth root of unity obtained by taking $k = 1$. Then

$$R = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}. \quad (8)$$

By De Moivre's theorem

$$R^k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \qquad (k\text{, an integer}).$$

Comparing with (7) we see that the nth roots of unity are powers of R. The n distinct roots of unity are

$$R, R^2, R^3, \cdots, R^{n-1}, R^n = 1. \quad (9)$$

Since the absolute value of R^k is 1, the points representing the nth roots of unity are equally spaced on the circumference of the unit circle. Joining these points by straight line segments a regular polygon of n sides is formed. The possibility of construction of such regular polygons with the use of straightedge and compass alone is taken up in a later chapter.

An nth root of unity which is not also a pth root $(p < n)$ is called a *primitive* root. The number R defined by (8) is a primitive nth root of unity.

Of the numbers (9), *those are primitive* nth *roots whose exponents are prime to* n.

Proof: Consider the root R^s in (9) $(s < n)$

$$R^s = \cos\frac{2s\pi}{n} + i\sin\frac{2s\pi}{n}.$$

Suppose s and n are not prime. Let k be the H.C.F. of s and n; then $n = ka$; $s = kb$; and $k = \frac{n}{a} < n$. We have

$$(R^s)^k = (R^s)^{\frac{n}{a}} = (\cos 2\pi + i\sin 2\pi)^{\frac{s}{a}} = 1^{\frac{s}{a}} = 1.$$

Since $k < n$, then R^s is not a primitive nth root of unity by virtue of the definition of a primitive nth root.

But if s and n have no common factor other than unity, then $(R^s)^r \neq 1$ for r a positive integer less than n. For

$$(R^s)^r = \cos \frac{2rs\pi}{n} + i \sin \frac{2rs\pi}{n}.$$

For this to be unity $\frac{rs}{n}$ must be an integer. However s is by hypothesis prime to n. Therefore $\frac{r}{n}$ must be an integer. But this is impossible, since $r < n$.

Exercises

Use De Moivre's theorem to solve the following binomial equations, and in each case plot all of the roots:

1. $x^2 - i = 0$
2. $x^2 + i = 0$
3. $x^4 + 1 = 0$
4. $x^5 - 1 = 0$
5. $x^5 + 1 = 0$
6. $x^6 - 1 = 0$
7. $x^3 - i = 0$
8. $x^3 + i = 0$
9. $x^4 - i = 0$
10. $x^4 + i = 0$
11. $x^8 - 1 = 0$
12. $x^6 + 1 = 0$

Evaluate the following:

13. $(\cos 30° + i \sin 30°)^{\frac{1}{3}}$
14. $(\cos 120° + i \sin 120°)^{\frac{1}{4}}$
15. $(\cos 150° + i \sin 150°)^{\frac{1}{5}}$
16. $(\cos 300° + i \sin 300°)^{\frac{1}{6}}$
17. $(1 + i)^{\frac{1}{3}}$
18. $(1 - i)^{\frac{1}{4}}$
19. $(-1 - i)^{\frac{1}{5}}$
20. $(1 + \sqrt{3i})^{\frac{1}{2}}$
21. $(-\sqrt{3} + i)^{\frac{1}{3}}$
22. $(1 - \sqrt{3}\, i)^{\frac{1}{3}}$
23. $\left(\frac{1}{2} - i\frac{\sqrt{3}}{2}\right)^{\frac{1}{2}}$
24. $\left(\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}\right)^{\frac{1}{3}}$
25. $\left(-\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}\right)^{\frac{1}{3}}$

26. Find the three cube roots of 27.
27. Find the four fourth roots of -81.
28. Find the five fifth roots of 32.
29. Find the six sixth roots of -64.
30. Find the six sixth roots of $+64$.
31. Show that the primitive nth roots of unity are for

(1) $n = 3: \omega, \omega^2$;
(2) $n = 4: i, -i$;
(3) $n = 5: R, R^2, R^3, R^4$.
(4) $n = 6: R, R^5$;
(5) $n = 8: R, R^3, R^5, R^7$.

32. Show that the product of the nth roots of unity is $+1$ if n is odd and -1 if n is even.

33. If R is an nth root of unity, prove that $\frac{1}{R}$ is also a root.

34. What are the values of $1 + \omega^n + \omega^{2n}$ for integral values of n?

35. What are the values of $i^n + i^{-n}$ for integral values of n?

36. Simplify $\omega^{13}, \omega^{23}, \omega^{-73}, \omega^{-48}, \omega^{-32}$.

37. Show that $0, 1 \pm i, 2$, are roots of the equation $(x - 1)^4 - 1 = 0$, and hence are vertices of a square.

CHAPTER IV

PROPERTIES OF POLYNOMIALS

4.1 Definitions. Whenever in mathematics one has two related variables, say x and y, and the relationship between them is such that for each value of x there is associated exactly one value of y, then y is said to be a one-valued function of x. This notion of function has been extended so as to apply to many-valued relations. In this course all functions encountered will be representable by formulas. A function is said to be *algebraic* if it may be represented by a formula which involves the variable through the operations of addition, subtraction, multiplication, division, involution, and evolution with constant rational exponents. All other functions are said to be *transcendental*. Thus,

$$3x^2 + ax - 5, \qquad \frac{2x - 3}{x - 2}, \qquad (3x^2 - x + 4)^{2/5},$$

are examples of algebraic functions; while

$\log x$, $\sin (x + y)$, 10^x, $\arccos x$, are examples of transcendental functions.

The reader may note that

$$\sin \left(\frac{\pi}{2} - \arcsin x\right)$$

although it might appear to be transcendental is actually algebraic, since, at least for restricted values of x, this reduces to $\pm\sqrt{1 - x^2}$.

An algebraic* function of x is an *integral rational function* if it may be represented by a polynomial in x with constant coefficients; namely, by an expression of the form

$$ax^n + bx^{n-1} + cx^{n-2} + \cdots + px + q. \tag{1}$$

* In this course we shall not need to consider functions which are not algebraic, or which, being algebraic, are not rational integral functions. Indeed we shall not pause to define rational functions or integral functions in general.

Here a, the coefficient of the term of highest degree in x, is called the leading coefficient. To avoid trivial cases we often require that the leading coefficients as written be different from 0. In such a case the polynomial is said to be properly of degree n.

A function of x is represented for brevity by $F(x)$, $f(x)$, $g(x)$, $\theta(x)$ or some other such symbol.

For certain values of x one polynomial may be equal to another differently constituted. The algebraic expression of such a relation is called an *equation.* Any value of x which, substituted for x in the equation, reduces it to an identity is said to be a *root* of the equation or zero of the polynomial. The determination of all possible roots constitutes the complete *solution of the equation.* By placing all terms of such an equation on the left-hand side of the equality sign and arranging in descending powers of x, we are able to put the equation in the form

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0, \quad (a_0 \neq 0). \quad (2)$$

The highest power of x in this equation being n, this equation is said to be an equation of the nth degree in x. If we divide each term by a_0, by hypotheses different from zero, the equation (2) can be put in the form

$$x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n = 0$$

where the coefficient of the term of highest degree is plus one. This is said to be the *p-form* of the equation. The term a_n (or p_n) which does not contain x, is called the *constant term.* An equation is numerical if its coefficients are numbers. An equation is said to be literal if its coefficients are letters which stand for numbers.

It is convenient to have special names for equations of different degrees.

A *linear* equation is of the first degree.

A *quadratic* equation is of the second degree.

A *cubic* equation is of the third degree.

A *quartic* equation is of the fourth degree.

A *quintic* equation is of the fifth degree.

A *sextic* equation is of the sixth degree.

4.2 Sign of a polynomial. Our search for the roots of an equation is facilitated if we know that our search should be restricted

to numbers less than some determinable positive number or greater than some determinable negative number; hence the desirability of the following theorem.

Theorem I: *If in the polynomial, with real coefficients,*

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n \quad (a_0 > 0)$$

the value $\left|\frac{a_k}{a_0}\right| + 1$, *or any greater value, be substituted for* x, *where* a_k *is that one of the coefficients* $a_1, a_2, \cdots, a_n$ *whose numerical value is greatest, the term containing the highest power of* x *will exceed, numerically, the sum of all the terms which follow.*

The inequality

$$a_0x^n > a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n$$

is satisfied by any positive value of x which makes

$$a_0x^n > |a_k|(x^{n-1} + x^{n-2} + \cdots + x + 1),$$

where a_k is the greatest, in numerical value, among the coefficients $a_1, a_2, \cdots, a_n$. Sum the geometric series within the parenthesis. We find

$$a_0x^n > |a_k|\frac{x^n - 1}{x - 1}, \quad \text{or} \quad x^n > \frac{|a_k|}{a_0(x-1)}(x^n - 1).$$

This inequality is satisfied if $a_0(x - 1) \geq |a_k|$; or, solving for x, if

$$x \geq \left|\frac{a_k}{a_0}\right| + 1, \quad \text{since} \quad a_0 > 0.$$

This theorem supplies us with a number, greater than unity, such that if x has this value or any greater value, the value of the polynomial will be positive.

If x is negative then a_0x^n will be positive or negative according as n is even or odd. By this method we obtain a negative value of x:

$$x_1 = -\left|\frac{a_k}{a_0}\right| - 1$$

such that for $x \leq x_1$ the term a_0x^n, for n even, will be greater than the sum of all the terms that follow it; while for n odd, this same term will be less than the sum of all the terms that follow it. In other words, the value of the polynomial for $x \leq x_1$ will be positive for n even, and negative for n odd.

In terms of roots of the equation obtained by equating the given polynomial to zero, this theorem means that *the equation cannot have a root greater than* $\left|\frac{a_k}{a_0}\right| + 1$ *or less than* $-\left|\frac{a_k}{a_0}\right| - 1$.

Example. Find an upper bound to the roots of

$$10x^3 - 37x^2 + 22x - 3 = 0.$$

In the test formula put $a_0 = 10$, $a_k = -37$. We have

$$x \geq \frac{37}{10} + 1 = 4.7.$$

We conclude that the given equation has no root greater than 4.7 or less than -4.7. The roots of the given equation are $\frac{1}{5}$, $\frac{1}{2}$, 3.

Theorem II: *If in the polynomial, with real coefficients,*

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \qquad (a_n \neq 0)$$

the value $\frac{|a_n|}{|a_n| + |a_k|}$, *or any smaller positive value, be substituted for* x, *where* a_k *is that one of the coefficients* $a_0, a_1, a_2, \cdots, a_{n-1}$ *whose numerical value is greatest, the term* a_n *will be numerically greater than the sum of all the others.*

In order to prove this, put $x = \frac{1}{y}$ and then multiply every term by y^n. We have

$$a_ny^n + a_{n-1}y^{n-1} + a_{n-2}y^{n-2} + \cdots + a_1y + a_0.$$

By Theorem I, if $y \geq \left|\frac{a_k}{a.}\right. + 1$, then

$$|a_n|\, y^n > a_{n-1}y^{n-1} + a_{n-2}y^{n-2} + \cdots + a_1y + a_0\,,$$

whence

$$|a_n| > a_{n-1}x + a_{n-2}x^2 + \cdots + a_1x^{n-1} + a_0x^n.$$

But if

$$y \geq \frac{|a_k|}{|a_n|} + 1, \quad \text{then} \quad x \leq \frac{|a_n|}{|a_n| + |a_k|},$$

which completes the proof of the theorem.

This theorem is often stated as follows:

Corollary I: Positive values so small may be assigned to x as to make the polynomial, with real coefficients, and without constant term,

$$a_0x^n + \cdots + a_{n-2}x^2 + a_{n-1}x$$

less than any preassigned positive quantity.

This statement of the theorem follows at once from the above proof by giving to a_n any positive value, however small. The following is another useful statement of the theorem:

Corollary II: When a sufficiently small positive value is assigned to x, the sign of the polynomial, with real coefficients,

$$a_0x^n + \cdots + a_{n-r}x^r \qquad (a_{n-r} \neq 0,\ n > r)$$

is the same as the sign of its term $a_{n-r}x^r$, of lowest degree.

This appears by writing the expression in the form

$$x^r(a_0x^{n-r} + \cdots + a_{n-r});$$

for when a sufficiently small positive value is given to x, the numerical value of a_{n-r} exceeds the sum of the other terms of the expression within the parenthesis, and the sign of that expression will consequently depend on the sign of a_{n-r}.

Example. Find a value of x such that for that value of x and for any smaller value the constant term will be numerically greater than the sum of all the other terms of

$$3x^4 + 6x^3 + 9x^2 + 4x - 1.$$

Put in the formula $|\,a_n\,| = 1$, $|\,a_k\,| = 9$. We have $x \leq \dfrac{1}{1+9} =$ 0.1. Substituting $x = 0.1$ in the given polynomial, we have

$$3x^4 + 6x^3 + 9x^2 + 4x = 0.4963 < 1.$$

Exercises

Find a value of x such that for that value of x or any larger value the term of highest degree will be greater than the sum of all of the other terms.

1. $x^3 - 6x^2 + 11x - 6$
2. $2x^3 - 5x^2 - 2x - 3$
3. $10x^3 - 22x^2 - 22x - 6$
4. $5x^7 - 5x^6 + 3x^5 + 10x^3 + x^2 - 1$
5. $2x^{17} + 3x^{11} - 6x^5 + 1$

6. $4x^3 - 10x^2 - 100x - 2000$
7. $4x^5 - x^4 + 2x^3 - x^2 - 4$
8. $100x^3 + x^2 + x + 1$
9. $x^{13} + 1$
10. $x^6 - 1$
11. $x^2 + x + 1$
12. $10x^3 + 18x^2 - 17x - 16$
13. $5x^4 + 10x^3 + 9x + 9$
14. $4x^5 + 3x^4 + 8x^2 + 7$
15. $3x^6 + 2x^4 + 9x^3 - 8x + 1$
16. $8x^3 + 2x^2 + x + 1$

Find a value of x such that for that value of x or any smaller positive value the constant term will be numerically greater than the sum of all the other terms.

17. $x^3 + x^2 + x + 1$
18. $x^2 + 3x - 4$
19. $x^5 - 4x^4 + 2x^3 + 1$
20. $x^5 - x^4 + 2x^3 - 2$
21. $x^5 + x^4 + x^3 + x^2 - 4$
22. $997x^3 + 100x^2 + 10x - 3$
23. $3x^4 + 5x^3 - 8$
24. $9998x^4 + 100x^2 - 2$
25. $996x^3 + 200x^2 + 10x - 4$
26. $x^3 + 99.99x^2 + 3x + 0.01$
27. $x^4 + 9.99x^3 + 3x^2 + x + 0.01$
28. $x^3 + 9.999x^2 + x - 0.001$
29. $8x^3 + 7x^2 + 9.9999x - 0.0001$
30. $9x^3 + 9x^2 + 9.999x - 0.001$
31. $9.99x^4 + 9.99x^3 + 9.99x - 0.01$
32. $98x^4 + 99x^3 + 97x^2 + 99x - 1$

Find a value of x such that for that value of x and for any smaller value the term of lowest degree will be numerically greater than the sum of all the other terms.

33. $3x^3 - 5x^2 + x$
34. $4x^4 + 3x^3 + 2x^2 - 1$
35. $9x^5 + 8x^4 - 3x^3 - x^2$
36. $98x^5 + 7x^4 - 5x^3 + 3x^2 - 2x$
37. $x^6 + 3x^5 - 97x^2 + 3x^3 - 3x^2$
38. $5x^6 - 3x^4 - 96x^3 + 2x^2 - 4x$
39. $2x^7 - 5x^4 - 95x^3 + 3x^2 + 5x$
40. $5x^6 - 6x^5 + 4x^4 + 3x^2$

4.3 Derived functions. Let us examine the form assumed by the polynomial when $x + h$ is substituted for x. Represent the polynomial by $f(x)$. Then

$$f(x) \equiv a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n.$$

Replacing x by $x + h$, we have

$$f(x + h) = a_0(x + h)^n + a_1(x + h)^{n-1} + a_2(x + h)^{n-2} + \cdots + a_{n-1}(x + h) + a_n.$$

By the binomial theorem

$$a_0(x + h)^n = a_0x^n + na_0x^{n-1}h + \frac{n(n-1)}{1\cdot 2}\cdot a_0x^{n-2}h^2 + \cdots + a_0h^n$$

$$a_1(x + h)^{n-1} = a_1x^{n-1} + (n-1)a_1x^{n-2}h + \frac{(n-1)(n-2)}{1\cdot 2}a_1x^{n-3}h^2 + \cdots + a_1h^{n-1}$$

$$a_2(x + h)^{n-2} = a_2x^{n-2} + (n-2)a_2x^{n-3}h + \frac{(n-2)(n-3)}{1\cdot 2}a_2x^{n-4}h^2 + \cdots + a_2h^{n-2}$$

$$a_{n-2}(x + h)^2 = a_{n-2}x^2 + 2a_{n-2}xh + a_{n-2}h^2$$

$$a_{n-1}(x + h) = a_{n-1}x + a_{n-1}h$$

$$a_n = a_n$$

Substitute these values in $f(x + h)$ and arrange in powers of h. We then have

$$f(x + h) = f(x) + f'(x)h + \frac{f''(x)}{1\cdot 2}h^2 + \cdots + \frac{f^{(n-1)}(x)}{1\cdot 2\cdots(n-1)}h^{n-1} + a_0h^n \qquad (3)$$

where

$$f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-2}x^2 + a_{n-1}x + a_n$$

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + (n-2)a_2x^{n-3} + \cdots + 2a_{n-2}x + a_{n-1}$$

$$f''(x) = n(n-1)a_0x^{n-2} + (n-1)(n-2)a_1x^{n-3} + (n-2)(n-3)a_2x^{n-4} + \cdots + 2a_{n-2}$$

Observe that a_0h^n may be written $\dfrac{f^{(n)}(x)}{1\cdot 2 \cdots n} h^n$.

The formula (3) is a case of what is known as *Taylor's theorem,* here stated, only for polynomials of degree n with real coefficients.

We observe that $f'(x)$ can be derived from $f(x)$ as follows: *Multiply each term in $f(x)$ by the exponent of x in that term, and decrease the exponent of x in that term by one. The sum of all terms obtained from $f(x)$ in this manner, is $f'(x)$.*

Observe that $f''(x)$ is obtained from $f'(x)$ in the same manner that $f'(x)$ was obtained from $f(x)$. In like manner the succeeding derivatives may all be obtained by successive operations of this character.

$f'(x)$ is called the first derivative of $f(x)$.
$f''(x)$ is called the second derivative of $f(x)$.
$f^{(r)}(x)$ is called the rth derivative of $f(x)$.

Obviously $f^{(r)}(x)$ is the first derivative of $f^{(r-1)}(x)$.

Example.

$$\begin{aligned}
f(x) &= x^5 + 2x^4 - 5x^3 - 3x^2 + 4x + 7 \\
f'(x) &= 5x^4 + 8x^3 - 15x^2 - 6x + 4 \\
f''(x) &= 20x^3 + 24x^2 - 30x - 6 \\
f'''(x) &= 60x^2 + 48x - 30 \\
f^{\text{IV}}(x) &= 120x + 48 \\
f^{\text{V}}(x) &= 120.
\end{aligned}$$

All succeeding derivatives are equal to zero.

Exercises

Find the first and second derivatives of

1. $x^4 + 3x^3 + 5x^2 - 1$
2. $x^5 - 7x^3 + 3x^2 - 8$
3. $3x^5 + 5x^4 - 2x^3 + 4x - 1$
4. $2x^4 + 3x^3 - 4x^2 + 5x - 7$
5. $x^3 - 6x^2 + 11x - 6$
6. $x^3 - x^2$
7. $x^3 - 7x^2 + 12x$
8. $x^4 - 5x^2 + 4$
9. $x^4 + 3x^3 + 4x^2 + 12x + 40$
10. $x^5 + 2x^4 - 6x^3 - 12x^2 - 15x - 30.$

4.4 Continuity of a polynomial. A polynomial $f(x)$, with real coefficients, is said to be continuous at $x = a$, where a is a real constant, if the difference

$$f(a + h) - f(a)$$

can be made numerically less than any preassigned positive number for all real values of h sufficiently small numerically.

By Taylor's theorem if $f(x)$ is a polynomial, then

$$f(a+h) - f(a) = f'(a)h + f''(a)\frac{h^2}{2} + \cdots + a_0 h^n.$$

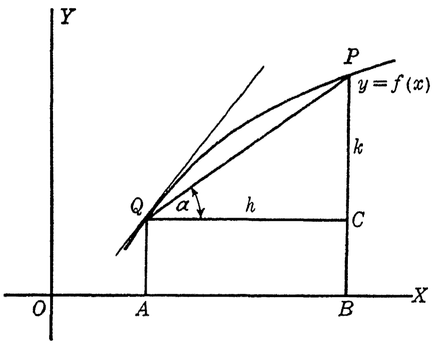


Fig. 11

By Cor. I, Theorem II, the right member of this equality can be made numerically less than any preassigned positive quantity; as a result $f(a+h) - f(a)$ can be made numerically as small as we please, and will ultimately vanish with h.

Hence a polynomial $f(x)$, with real coefficients, is a continuous function of x at every real value of x.

4.5 Geometric interpretation of a derivative.

Let $y = f(x)$ be represented by a continuous smooth curve QP. Let Q be a point on the curve whose coordinates are (x, y).

Let P be a point on the curve whose coordinates are $(x+h, y+k)$. Then

$$y = f(x) = AQ,$$

$$y + k = f(x+h) = BP,$$

whence

$$k = f(x+h) - f(x) = CP.$$

Then

$$\tan\alpha = \frac{f(x+h) - f(x)}{h} = \frac{CP}{QC}.$$

By Taylor's theorem

$$f(x+h) - f(x) = f'(x)h + f''(x)\frac{h^2}{2} + \cdots + f^{(n)}(x)\frac{h^n}{n}$$

whence

$$\frac{f(x+h)-f(x)}{h} = f'(x) + h\left\{\frac{1}{2}f''(x) + \cdots + f^{(n)}(x)\,\frac{h^{n-2}}{\lfloor n}\right\}.$$

Hence, the limit of $\dfrac{f(x+h)-f(x)}{h}$ as h approaches zero is $f'(x)$. But as h approaches zero, k also approaches zero, and the point P approaches Q, while the secant line PQ approaches as a limiting position the tangent line to the curve at Q. Thus we see that *$f'(x)$ represents the slope* of the tangent line to the curve $y = f(x)$ at the point (x, y).*

Example. Find the slope of $y = x^3 - x$ at $x = 1$, $y = 0$. $f(x) = x^3 - x; f'(x) = 3x^2 - 1$. For $x = 1$, $3x^2 - 1 = 2$. Therefore the slope of $y = x^3 - x$ at $x = 1$, is $+2$.

Exercises

Find the slope of the given curve at the point stated.

1. $y = x^3 - 3x;\ x = 0$
2. $y = 2x^3 - 15x^2 + 36x + 23;\ x = 1$
3. $y = x^2 - 5x + 6;\ x = 3$
4. $y = x^4 - 1;\ x = 2$
5. $y = x^4 - 8x^3 - 7x;\ x = 3$
6. $y = x^5 - 3x^2 + 5x;\ x = 2$
7. $y = x^4 - 3x^3 + 4x^2 - 7;\ x = 2$
8. $y = x^6 + 3x^2 - 9x;\ x = 1$
9. $y = x^3 - 6x^2 + 11x - 6;\ x = 2$
10. $y = x^3 - 4x^2 + 3x - 5;\ x = 3$
11. $y = x^4 - 5x^3 + 6x^2 - 10;\ x = 2$
12. $y = x^4 - 6x^3 + 4x^2 - 8;\ x = 1$
13. $y = 2x^4 - 3x^3 + 4x^2 - 7x;\ x = -2$
14. $y = x^5 - 3x^3 + 7x - 8;\ x = -1$
15. $y = 3x^4 - 7x^3 + 5x^2 - 4;\ x = -2$
16. $y = 2x^5 - 5x^4 + 3x^2 - 10;\ x = -1$
17. In each of the above exercises find the second derivative.

4.6 The Remainder Theorem. Consider the operation of dividing a polynomial of degree n, say

$$A(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \quad (a_0 \neq 0),$$

* By slope of the tangent line is meant the tangent of the angle that the tangent line makes with the positive direction of the x-axis.

by a binomial of the form $x - h$. On division, we have

$$A(x) \equiv (x - h)Q(x) + R,$$

where $Q(x)$, the quotient, is a polynomial of degree $n - 1$, and R, the remainder, is a constant, in some cases zero. This is an identity in x. For $x = h$, since $x - h$ is then zero and $Q(h)$ is determinate and finite, the later relation reduces to $A(h) = R$. Hence we have for any polynomial, $f(x)$, the

Remainder Theorem. *If a polynomial $f(x)$ be divided by the binomial $x - h$, the remainder R is the value of $f(x)$ when x takes on the value h.*

To illustrate a use for the remainder theorem consider $x(x-1)(x-2)(x-3) - 24$. At first sight it is not obvious that this polynomial will be divisible by $x - 4$. However, if x be replaced by 4, the value of the polynomial is zero; hence the remainder on dividing by $x - 4$ is zero. Similarly it is divisible by $x + 1$. Since the polynomial is divisible by both $x - 4$ and $x + 1$, it is divisible by their product $x^2 - 3x - 4$. Upon expansion we have

$$x(x - 1)(x - 2)(x - 3) - 24 = x^4 - 6x^3 + 11x^2 - 6x - 24.$$

On dividing this polynomial by $x^2 - 3x - 4$, we find that the division is exact, with quotient $x^2 - 3x + 6$.

The remainder theorem may be stated as follows:

The value of a polynomial $f(x)$ when h is substituted for x is equal to the remainder when $f(x)$ is divided by $x - h$.

If $R = 0$, the division is exact. Hence we have the following Factor Theorem: *If h is a root of the equation $f(x) = 0$, then $x - h$ is a factor of $f(x)$; and, conversely, if $x - h$ is a factor of $f(x)$, h is a root of $f(x) = 0$.*

Example 1: Let us divide $x^3 - 5x^2 + 4x + 7$ by $x - 2$. We have

$$x^3 - 5x^2 + 4x + 7 = (x - 2)(x^2 - 3x - 2) + 3.$$
$$f(2) = 2^3 - 5\cdot 2^2 + 4\cdot 2 + 7 = 3,$$

which is the remainder.

Example 2: $x = 3$ is a root of $f(x) \equiv x^3 - 2x^2 + x - 12 = 0$, for $f(3) = 3^3 - 2\cdot 3^2 + 3 - 12 = 0$. Then $x - 3$ is a factor of $x^3 - 2x^2 + x - 12$. We have

$$x^3 - 2x^2 + x - 12 = (x - 3)(x^2 + x + 4).$$

Exercises

Without performing the division find the remainder when

1. $x^3 - 4x^2 + 3x + 5$ is divided by $x - 1$.
2. $2x^3 + 3x^2 - 4x - 8$ is divided by $x - 2$.
3. $x^4 + 5x^3 - 4x + 20$ is divided by $x + 2$.
4. $x^5 - 4x^4 + 3x^3 + 5x - 10$ is divided by $x - 3$.
5. $x^4 + 2x^3 + x^2 - x + 5$ is divided by $x + 2$.
6. $2x^6 - 5x^4 + 4x^3 + 3x^2 - 6x + 3$ is divided by $x + 1$.
7. $3x^8 - 7x^5 + 5x^3 + 2x + 4$ is divided by $x + 1$.
8. $2x^5 - 3x^4 + 4x^3 - 7x - 24$ is divided by $x - 2$.

Without performing the division show that

9. $x^3 - 6x^2 + 11x - 6$ is divisible by $x - 1$, $x - 2$, $x - 3$.
10. $x^3 + 4x^2 - 9x - 36$ is divisible by $x - 3$, $x + 3$, $x + 4$.
11. $x^3 + x^2 - 14x - 24$ is divisible by $x + 2$, $x + 3$, $x - 4$.
12. $x^4 - 8x^3 + 17x^2 + 2x - 24$ is divisible by $x - 2$, $x - 3$, $x + 1$, $x - 4$.
13. $x^4 + 6x^3 + 12x^2 + 11x + 6$ is divisible by $x + 2$, $x + 3$.
14. Show that $x^n - a^n$ is divisible by $x - a$.
15. Show that $x^n + a^n$ is divisible by $x + a$ if n is odd.
16. Show that $x^n - a^n$ is divisible by $x + a$ if n is even.

4.7 Synthetic division. When plotting the graph $y = f(x)$, the fact expressed by the Remainder Theorem is frequently used to obtain the values of $f(x)$ for convenient choices of x.

Suppose we are interested in graphing $y = f(x)$, where

$$y = 2x^5 - 30x^3 + 2x^2 + 65x - 7.$$

It is clear that for $x = 0$, $y = -7$, but for other integral values of x at least a little computation is necessary. Synthetic division gives a short and convenient method of making these computations. Divide

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \tag{4}$$

by $x - h$. Let $Q(x)$ be the quotient and the remainder R. $Q(x)$ will be of degree $n - 1$ in x. Put

$$Q(x) \equiv b_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-1}.$$

Multiply $Q(x)$ by $x - h$, add R, and collect coefficients of like powers of x. We have

$$\begin{aligned} f(x) = (x - h)Q(x) + R = b_0x^n &+ (b_1 - hb_0)x^{n-1} \\ + (b_2 - hb_1)x^{n-2} + \cdots &+ (b_{n-1} - hb_{n-2})x + R - hb_{n-1}. \end{aligned} \tag{5}$$

Coefficients of like powers of x in (4) and (5) are equal. Then

$$b_0 = a_0, \qquad b_1 = a_1 + hb_0,$$

$$b_2 = a_2 + hb_1, \cdots, b_{n-1} = a_{n-1} + hb_{n-2}, \qquad R = a_n + hb_{n-1}.$$

The coefficients $b_0, b_1, \cdots, b_{n-1}$ and the remainder R can be readily computed successively by means of these equations. The work is facilitated by the following scheme of tabulation:

$$\begin{array}{cccccc} a_0 & a_1 & a_2 & a_3 \cdots & a_{n-1} & a_n \\ & hb_0 & hb_1 & hb_2 \cdots & hb_{n-2} & hb_{n-1} \\ \hline a_0 = b_0 & b_1 & b_2 & b_3 \cdots & b_{n-1} & R \end{array}$$

Rule for synthetic division. *To divide $f(x)$ by $x - h$, arrange $f(x)$ in descending powers of x. Supply missing powers of x by the proper power of x with zero as a coefficient.*

Write the coefficients in a horizontal line in the order $a_0, a_1, \cdots, a_n$. In the second space below a_0, write a_0 (call it b_0). Multiply b_0 by h and add the product hb_0 to a_1. Write the sum b_1 in the second space below a_1. Multiply b_1 by h, add to a_2, and write the sum b_2 in the second space below a_2. Continue this process. The last sum is the remainder R. The sums $b_0, b_1, \cdots, b_{n-1}$ are the coefficients of the powers of x in the quotient, $Q(x)$, arranged in descending order.

Example. Divide $x^4 + 2x^3 - 5x^2 - 7$ by $x - 2$.

$$\begin{array}{l|l} 1 + 2 - 5 + 0 - 7 & \underline{2} \\ + 2 + 8 + 6 + 12 & \\ \hline 1 + 4 + 3 + 6 + 5 & \end{array}$$

The quotient is $x^3 + 4x^2 + 3x + 6$. The remainder is 5.

Synthetic division is much used. It is important that it be thoroughly understood by the student. To this end the division in the above example will be discussed in more detail.

By long division we have

$$\begin{array}{l} x^4 + 2x^3 - 5x^2 + 0\cdot x - 7 \,|\, x - 2 \\ \underline{x^4 - 2x^3} \qquad\qquad\qquad\qquad x^3 + 4x^2 + 3x + 6 = Q(x) \\ \quad + 4x^3 - 5x^2 \\ \quad \underline{+ 4x^3 - 8x^2} \\ \qquad\qquad + 3x^2 + 0\cdot x \\ \qquad\qquad \underline{+ 3x^2 - 6x} \\ \qquad\qquad\qquad + 6x - 7 \\ \qquad\qquad\qquad \underline{+ 6x - 12} \\ \qquad\qquad\qquad\qquad + 5 = R \end{array}$$

Since like powers of x are written in the same vertical column, there can be no confusion if only the coefficients are written. The work will then appear as follows:

$$
\begin{array}{llllll|l}
1 & +2 & -5 & +0 & -7 & & \underline{1-2} \\
\underline{1} & \underline{-2} & & & & & 1+4+3+6 \\
 & +4 & -5 & & & & \\
 & \underline{+4} & \underline{-8} & & & & \\
 & & +3 & +0 & & & \\
 & & \underline{+3} & \underline{-6} & & & \\
 & & & +6 & -7 & & \\
 & & & \underline{+6} & \underline{-12} & & \\
 & & & & +5 & &
\end{array}
$$

The first term in the divisor may be omitted, since all divisors have the same first term. Only the first term of each partial remainder need be written down, for the second term is a repetition of the term directly above it in the dividend. We observe that it is not necessary to write the coefficients of the quotient, for these are equal respectively to the first coefficients of the dividend and the successive remainders. The coefficients of the first terms of the successive products may also be omitted. The work will then appear as follows:

$$
\begin{array}{lllll|l}
1 & +2 & -5 & +0 & -7 & \underline{-2} \\
 & -2 & & & & \\
 & +4 & & & & \\
 & & \underline{-8} & & & \\
 & & +3 & & & \\
 & & & \underline{-6} & & \\
 & & & +6 & & \\
 & & & & \underline{+12} & \\
 & & & & +5 &
\end{array}
$$

Obviously this can be written more compactly in this manner:

$$
\begin{array}{lllll|l}
1 & +2 & -5 & +0 & -7 & \underline{-2} \\
 & -2 & -8 & -6 & -12 & \\
\hline
1 & +4 & +3 & +6 & +5 &
\end{array}
$$

If we replace -2 by $+2$, and *add* the partial products instead of *subtracting* them, we obtain the same result. Our work now appears as follows:

$$\begin{array}{rrrrr|l} 1 & +2 & -5 & +0 & -7 & \underline{2} \\ & +2 & +8 & +6 & +12 & \\ \hline 1 & +4 & +3 & +6 & +5 & \end{array}$$

Even this second row of numbers may be carried mentally, in which case the work would appear as shown below:

$$\begin{array}{rrrrr|l} 1 & +2 & -5 & +0 & -7 & \underline{2} \\ \hline 1 & +4 & +3 & +6 & +5 & \end{array}$$

Exercises

In the following find the quotient and remainder. Use synthetic division.

1. Divide $x^4 - 3x^3 - 7x^2 + 5x + 3$ by $x - 2$.
2. Divide $x^5 + 4x^4 - 5x^3 + 2x^2 - 3x + 5$ by $x + 2$.
3. Divide $x^5 - 5x^3 + 3x^2 - 7$ by $x - 3$.
4. Divide $2x^3 - 4x^2 - 17x + 8$ by $x - 4$.
5. Divide $2x^4 - 5x^3 - 4x^2 + 3x - 7$ by $x - 3$.
6. Divide $2x^5 - 3x^4 - 8x^3 + 5x - 10$ by $x - 2$.
7. Divide $3x^4 - 8x^3 - 2x^2 - x + 3$ by $x - 3$.
8. Divide $2x^4 + 5x^3 + 3x^2 + 4x + 5$ by $x + 2$.
9. Divide $x^5 + 32$ by $x + 2$.
10. Divide $x^6 - 64$ by $x - 2$.
11. Divide $3x^4 + 10x^3 - 7x^2 - 6x - 7$ by $x + 4$.
12. Divide $x^5 - 2x^4 + x^3 - 3x^2 + 5$ by $x - 2$.

4.8 Graphic representation.

The roots of the equation

$$x^2 - 8x + 9 = 0$$

can be found graphically as follows: Put

$$y = x^2 - 8x + 9.$$

By synthetic division, or otherwise, find and tabulate values of y for integral values of x differing by unity. We have

x	0	1	2	3	4	5	6	7	8
y	9	2	-3	-6	-7	-6	-3	2	9

Plot the corresponding points and join by as smooth a curve as possible. The x coordinates of the points where the curve crosses the x-axis are the roots of the given equation. Measuring, we find 1.3 and 6.6 to be the roots of the given equation.

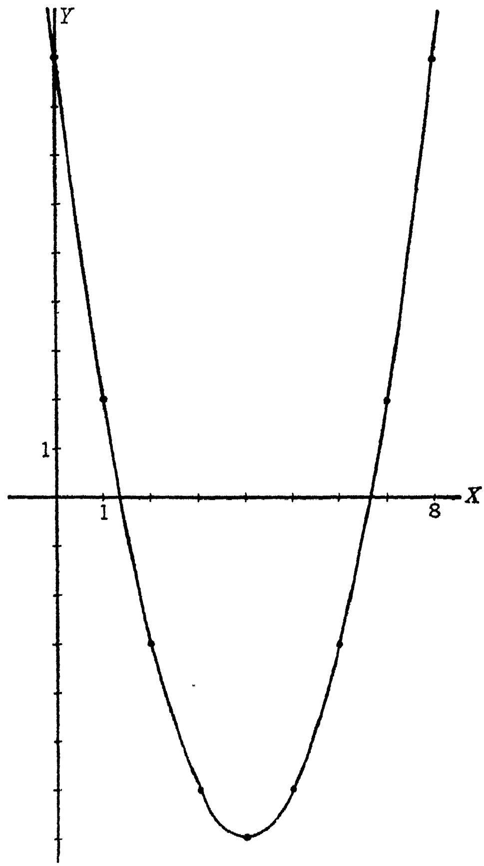


Fig. 12

Using this method to find the roots of

$$x^2 - 5x + 6 = 0$$

put

$$y = x^2 - 5x + 6.$$

Tabulating values of y for integral values of x, we have

x	0	1	2	3	4	5
y	6	2	0	0	2	6

We see at once from the table that $x = 2$ and $x = 3$ are the roots, since for these values of x, $y = 0$.

The following example will make it clear that caution must be used in plotting a curve. It does not always suffice to assign successive integral values to x. For example, to find the roots of

$$6x^3 - 17x^2 + 11x - 2 = 0,$$

put

$$y = 6x^3 - 17x^2 + 11x - 2.$$

Assign values to x, compute y and tabulate the corresponding values. We have

x	-1	0	1	2	3
y	-36	-2	-2	0	40

Plotting these points, all but the first and last of which are represented by large black dots in the accompanying figure 13, we might be led to think that the curve crossed the x-axis only once, namely, at $x = 2$. As a matter of fact the curve crosses the x-axis also at $x = \frac{1}{2}$ and $x = \frac{1}{3}$. In plotting this curve, it becomes necessary to assign other than integral values to x. As one may see, there is a possibility of serious error in plotting a curve, no matter how close together are the values assigned to x. It is essential that one locate all points like A and B (Fig. 13) where the curve changes direction. This we will learn to do in the next article.

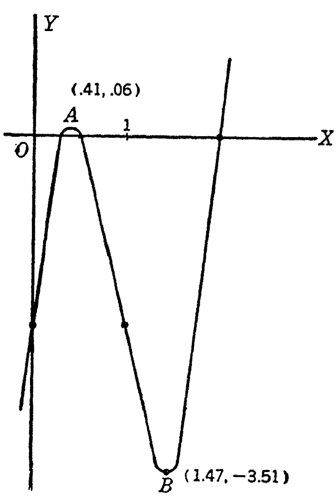


Fig. 13

4.9 Maxima and minima. A point like B, in the neighborhood of which there are no lower points on the curve, is called a *mini-*

mum point. A point like A, in the neighborhood of which there is no higher point on the curve is called a *maximum* point. At every maximum and minimum point on a curve which represents so smooth a curve as a polynomial it is clear that the tangent line is parallel to the x-axis, and therefore the slope of the tangent line is zero. But $f'(x)$ represents the slope of the tangent line. Therefore at every maximum and minimum point on a polynomial curve $f'(x) = 0$. In order to find the maximum and minimum points on such a curve, find those values of x for which $f'(x) = 0$. Substitute these values of x in $y = f(x)$. If there are any maximum or minimum points on the curve, their coordinates will be found among the pairs of values of x and y so obtained.

A point of horizontal inflection is a point like P or Q (Fig. 14) at which the curve crosses its tangent line, the tangent line being

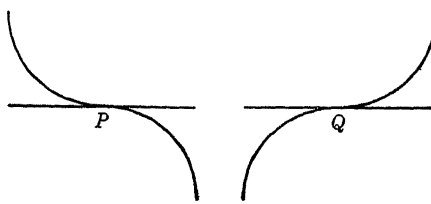


Fig. 14

parallel to the x-axis. At such a point also $f'(x) = 0$. In the simpler cases, the student will have no difficulty in determining whether the point under consideration is a maximum, minimum, or a point of horizontal inflection.

We give here, in mathematical terms, the conditions for a maximum, minimum, or point of horizontal inflection at $x = x_1$

If there exists a region $|x - x_1| < \delta$, $\delta > 0$, about x_1 such that for every positive $h < \delta$, (Fig. 15),

(a) $f(x_1 \pm h) > f(x_1)$, then there is a minimum at $x = x_1$;
(b) $f(x_1 \pm h) < f(x_1)$, then there is a maximum at $x = x_1$;
(c) either $f(x_1 - h) < f(x_1) < f(x_1 + h)$ and $f'(x_1) = 0$ [Point Q],
or $f(x_1 - h) > f(x_1) > f(x_1 + h)$ and $f'(x_1) = 0$ [Point P],

then there is a point of horizontal inflection at $x = x_1$.

Example: By the help of the graphical method find the roots of $6x^3 - 35x^2 + 64x - 35 = 0$. Put

$$y = 6x^3 - 35x^2 + 64x - 35.$$

Tabulate values of x and y.

x	0	1	2	3
y	-35	0	1	4

$$f(x) = 6x^3 - 35x^2 + 64x - 35$$

$$f'(x) = 18x^2 - 70x + 64$$

$$f'(x) = 0 \text{ gives } x = 1.47; 2.4\dot{2}$$

Whence $A(1.47, 2.45)$ is a maximum point
and $B(2.42, -0.06)$ is a minimum point.

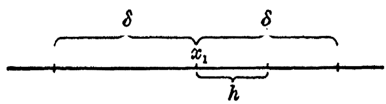


Fig. 15

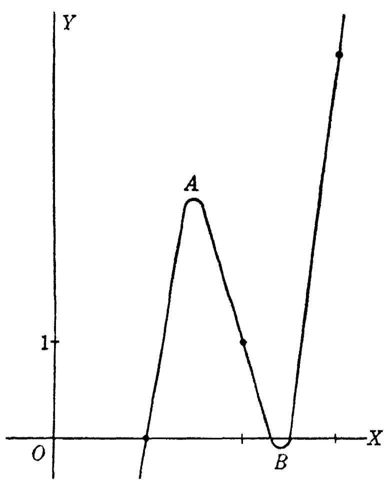


Fig. 16

The curve can now be drawn with reasonable accuracy. The roots are $x = 1, 5/2, 7/3$.

If the example had been a little different so that the minimum point B were above the x-axis, then, obviously, there could be but one real root. Thus

$$6x^3 - 34x^2 + 63x - 35 = 0$$

or

$$(x - 1)(6x^2 - 28x + 35) = 0$$

has but one real root.

Exercises

Plot the following curves. Determine the maxima and minima.

1. $y = x^2 - 1$
2. $y = x^2 - 2$
3. $y = x^2 - 3x + 4$
4. $y = x^2 - 3x + 3$
5. $y = x^2 - 3x + 2$
6. $y = x^2 - 3x + 1$
7. $y = x^2 - 3x$
8. $y = x^2 - 3x - 1$
9. $y = x^3$
10. $y = x^3 - 3x$
11. $y = x(x - 1)(x - 4)$
12. $y = (x - 1)(x - 4)(x - 6)$
13. $y = (x + 1)(x - 1)(x - 4)$
14. $y = x^2(x - 1)(x - 4)$
15. $y = x^3(x - 4)$
16. $y = x^4$
17. $y = (x - 1)^2(x - 4)^2$
18. $y = x^2(x - 4)$
19. $y = (x - 4)^2(x + 2)$
20. $y = x^5$
21. $y = x^4(x - 5)$
22. $y = x^3(x - 3)^2$
23. $y = x^2(x - 1)(x - 5)$
24. $y = x^4 - 8x^2 + 8$
25. $y \quad x^3 - 6x^2 + 9x + 5$
26. $y \quad 2x^3 - 3x^2 - 36x + 60$
27. $y = x^4 - 8x^2 - 4$
28. $y = x^4 - 8x^2 + 1$
29. $y = x^4 - 8x^2 + 2$
30. $y = x^4 - 8x^2 + 16$
31. $y = x^4 - 8x^2 - 9$
32. $y = 6x^3 - 23x^2 + 16x - 3 = (x - 3)(2x - 1)(3x - 1)$

Find the roots of

33. $x^3 - 3x^2 + 4 = 0$
34. $x^3 + x^2 - 8x - 12 = 0$
35. $x^3 - x^2 - 8x + 12 = 0$
36. $x^3 - 4x^2 - 3x + 18 = 0$
37. $x^3 - 8x^2 + 5x + 50 = 0$
38. $x^3 + 5x^2 - 8x - 48 = 0$
39. $x^4 - 2x^3 - 12x^2 + 40x - 32 = 0$

Chapter V

ELEMENTARY THEOREMS ON THE ROOTS OF AN EQUATION

5.1 Roots between a and b, if f(a) and f(b) have opposite signs.

Theorem I: *If the coefficients of a polynomial $f(x)$ are real numbers, and if a and b are real numbers such that $f(a)$ and $f(b)$ have contrary signs, then the equation $f(x) = 0$ has at least one real root between a and b.*

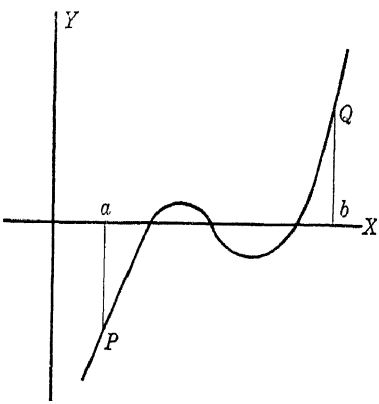


Fig. 17

No proof of this statement is given here. $f(x)$ is a continuous function of x. For every value of x between a and b there is one and only one point on the curve. Geometrically it is obvious that a continuous curve passing through P and Q on opposite sides of the x-axis must cross the x-axis at least once; and if the curve crosses more than once (but does not coincide with the axis throughout any segment), it must cross an odd number of times between $x = a$ and $x = b$.

5.2 The Factor theorem.

Let us quote here the factor theorem §4.6 of the preceding chapter: *If h is a root of the equation $f(x) = 0$, then $x - h$ is a factor of $f(x)$; and, conversely, if $x - h$ is a factor of $f(x)$, h is a root of $f(x) = 0$.*

Theorem II: *If $f(x)$ is a polynomial of degree n in x, then $f(x) = 0$ has n* roots and no more.*

We assume** that every rational, integral algebraic equation

* This does not necessarily mean n distinct roots.

** This assumption is justified, but the proof lies outside the scope of this volume.

has at least one root, real or imaginary. Put

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \quad (a_0 \neq 0).$$

Let r_1 be a root of $f(x) = 0$. By the factor theorem $f(x) = (x - r_1)Q_1(x)$ where $Q_1(x)$ is a polynomial of degree $n - 1$ with a_0 as leading coefficient. By our assumption $Q_1(x) = 0$ has a root. Call this root r_2. Then by the factor theorem

$$Q_1(x) = (x - r_2)Q_2(x)$$

and $$f(x) = (x - r_1)(x - r_2)Q_2(x).$$

$Q_2(x)$ is a polynomial of degree $n - 2$, again with a_0 as leading coefficient. $Q_2(x) = 0$ has a root. Continuing in this way we obtain n linear factors of $f(x)$ namely $x - r_1, x - r_2, \cdots, x - r_n$ and in addition the constant factor a_0, which is the coefficient of x^n in $f(x)$. Then

$$f(x) = a_0(x - r_1)(x - r_2) \cdots (x - r_n).$$

If x is given any one of the values $r_1, r_2, \cdots, r_n$, then $f(x)$ vanishes; so, $f(x)$ has for roots the n numbers $r_1, r_2, \cdots, r_n$. If x is given any value other than $r_1, r_2, \cdots, r_n$, then no factor of $f(x)$ can vanish, and the equation $f(x) = 0$ is not satisfied. Hence $f(x) = 0$ cannot have more than n roots.

From this theorem it follows that the problem of solving an equation $f(x) = 0$ is essentially the same as that of factoring the polynomial $f(x)$. In order to form the equation that has given numbers for its roots, we have merely to equate to zero the product of the differences obtained by subtracting each given number in turn from x.

Example 1. Form the equation of lowest degree whose roots are $-2, \frac{1}{2}, 1, 0$. We have

$$(x + 2)(x - \tfrac{1}{2})(x - 1)(x - 0) = 0;$$

the desired equation (when written with $a_0 = 2$) is

$$x(x - 1)(x + 2)(2x - 1) = 0, \quad \text{or} \quad 2x^4 + x^3 - 5x^2 + 2x = 0.$$

If any root is repeated, that root is written down as many times as it occurs as a root.

Example 2. Form the equation of lowest degree whose roots

are 1, 1, 1, 2. We have

$$(x-1)(x-1)(x-1)(x-2) = (x-1)^3(x-2)$$
$$= x^4 - 5x^3 + 9x^2 - 7x + 2 = 0.$$

We are now able to solve the following problem:

Problem. It is known that 1, 2 are roots of the equation $f(x) \equiv x^4 + x^3 - 7x^2 - x + 6 = 0$. Find the other roots. The product $(x-1)(x-2) = x^2 - 3x + 2$ is a factor of $f(x) = 0$. Factoring we find

$$x^4 + x^3 - 7x^2 - x + 6 = (x^2 - 3x + 2)(x^2 + 4x + 3) = 0.$$

The desired roots are the roots of $x^2 + 4x + 3 = (x+1)(x+3) = 0$, namely -1 and -3.

Exercises

Form the equations, in each case of lowest degree, whose roots are as follows:

1. $3, -2, -1$
2. $\frac{1}{2}, \frac{1}{3}, 2$
3. $-2, -3, 2$
4. $1, 2, 2, 3$
5. $0, 2, 3, 4$
6. $2, 2, 2, -3$
7. $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$
8. $\frac{2}{3}, \frac{3}{2}, 2$
9. $\frac{2}{3}, \frac{3}{4}, \frac{1}{2}$
10. $-\frac{2}{3}, -\frac{5}{3}, 3$
11. $3, \pm\sqrt{5}$
12. $\pm\sqrt{2}, \pm\sqrt{3}$
13. $5, 3 \pm \sqrt{2}$
14. $-1, 2 \pm \sqrt{3}$
15. $2 \pm \sqrt{3}, 3 \pm \sqrt{2}$
16. $3, \pm 2i$
17. $5, \pm 3i$
18. $4, 3 \pm 2i$
19. $2 \pm 3i, 3 \pm 2i$
20. $1 \pm 2i, 3 \pm \sqrt{2}$

5.3 Imaginary roots occur in pairs. Let us remind the reader that the number of real roots of an equation with real coefficients may be less than the degree of the equation. For example $x^4 - 2x^2 - 8 = (x^2 + 2)(x^2 - 4) = 0$ has two real roots, 2, -2 and two imaginary roots $\pm i\sqrt{2}$. Plotting $y = x^4 - 2x^2 - 8$ (Fig. 18) we find that the curve cuts the x-axis in two real points. If the equation had four real roots the curve would cut the x-axis in four real points.

Theorem III: *If $f(x)$ is a polynomial with real coefficients, and if $a + bi$, where a and b are real, and $b \neq 0$, is a root of the equation $f(x) = 0$, then $a - bi$ is also a root.*

Let

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n.$$

Putting $a + bi$ in place of x, we have

$$a_0(a + bi)^n + a_1(a + bi)^{n-1} + \cdots + a_{n-1}(a + bi) + a_n.$$

Expand each binomial and simplify. All terms that contain even powers of i will be real; all terms that contain odd powers of i will be pure imaginary. Represent by P the algebraic sum of all

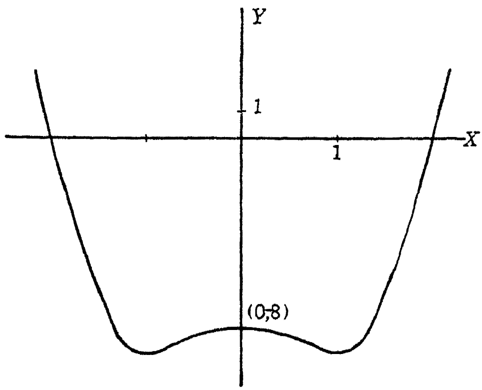


Fig. 18

of the real terms and by Qi the sum of all of the pure imaginary terms. Then, since by hypothesis $a + bi$ is a root of $f(x) = 0$, we have

$$f(a + bi) = P + Qi = 0.$$

Whence

$$P = 0 \text{ and } Q = 0, \quad \text{by} \quad §3.3$$

For any positive integral exponent the binomial expansion of $a - bi$ can be obtained from that of $a + bi$ by replacing i by $-i$. This replacement will not alter those terms in which bi is raised to an even power. In those terms in which bi is raised to an odd power, this replacement will result in a change of sign of the term, and we shall have

$$f(a - bi) = P - Qi.$$

But

$$P = 0 \quad \text{and} \quad Q = 0;$$

accordingly

$$f(a - bi) = P - Qi = 0,$$

and so

$$a - bi \text{ is a root of } f(x) = 0.$$

Corollary I. *Every polynomial equation $f(x) = 0$, with real coefficients, is the product of real factors, each of the first or second degree.*

For to each real root r, there corresponds the factor $x - r$; and to each pair of imaginary roots $a + bi$, $a - bi$, there corresponds the real quadratic factor $x^2 - 2ax + (a^2 + b^2)$, which is the product of $x - a - bi$ and $x - a + bi$.

Corollary II. *Every polynomial equation $f(x) = 0$, with real coefficients, of odd degree, has at least one real root.*

Since imaginary roots occur in pairs, the total number of imaginary roots must be even. For an equation of odd degree the total number of roots, real and imaginary, must be odd, and at least one root must be real. For a cubic equation, with real coefficients, this means that either all roots are real, or one root is real and two imaginary.

Illustration: The roots of $x^3 - 7x^2 + 25x - 39 = 0$ are 3, $2 \pm 3i$ and $x^3 - 7x^2 + 25x - 39 = (x - 3)(x^2 - 4x + 13)$.

5.4 Theorem IV. *Every polynomial equation $f(x) = 0$, with real coefficients, of even degree and with a negative constant term and coefficient of highest degree term positive, has at least one real positive root, and one real negative root.*

If $x = 0$, $f(x)$ is negative, since the constant term is negative. If $x = a$ is large enough, $f(a)$ is positive [§4.2]. Then there is a real root, which must be positive, between 0 and a. If $x = -b$, b real and positive, and b is large enough, $f(-b)$ is positive, since $f(x)$ is of even degree; and there is a real root, which must be negative, between 0 and $-b$ [§4.2].

Illustration: $x^4 - 3x^3 - 7x^2 + 27x - 18 = 0$ has for roots 1, 2, 3, -3.

With the exception of those equations of even degree with positive constant term, this theorem, with corollary II of the previous theorem, proves that every rational integral algebraic equation with real coefficients has a real root.

The following statements respecting the absence of real roots will be seen to be obviously true:

Statement I: If the coefficients in $f(x)$ are all positive or zero (not all being zero), the equation $f(x) = 0$ has no positive roots.

Statement II: If all the nonvanishing coefficients of the *even* powers of x in $f(x)$ have one sign, and all the nonvanishing coefficients of the *odd* powers of x the contrary sign, the equation $f(x) = 0$ has no negative root.

Illustration: $(x - 1)(x - 2)(x - 3) = x^3 - 6x^2 + 11x - 6 = 0$ has no negative root.

Statement III: If $f(x)$ has a term independent of x and involves only *even* powers of x and the nonvanishing coefficients are all of the same sign, the equation $f(x) = 0$ has no real roots.

Illustration: $x^4 + x^2 + 1 = 0$ has no real roots. The roots are $\pm\sqrt{\frac{1}{2}(-1 \pm i\sqrt{3})}$.

Statement IV: If $f(x)$, not identically zero, involves only *odd* powers of x and the nonvanishing coefficients are all of the same sign, the equation $f(x) = 0$ has no real root, except $x = 0$.

Illustration:

$$x^5 + 4x^3 + 4x = x(x^2 + 2)^2 = 0.$$

5.5 Theorem V: *If the polynomial $f(x)$, with terms arranged in descending order, consists of a set of terms in which the coefficients are all of one sign, followed by a set of terms in which the coefficients are all of the contrary sign, the equation $f(x) = 0$ has one and only one positive root.*

Let

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \quad (a_0 \neq 0).$$

Let us assume, for example, that the constant term is negative. Then $f(0) < 0$. For x large enough $f(x)$ has the same sign as a_0x^n which is positive. Then by theorem I, $f(x) = 0$ has at least one positive root. We shall now show, under the hypothesis of the theorem, that $f(x) = 0$ cannot have more than one positive root.

Suppose that $a_0, a_1, \cdots, a_r$ are positive or zero (not all being zero) and the remaining coefficients (not all zero) are negative or zero.

Let

$$a_{r+1} = -P_{r+1}, \qquad a_{r+2} = -P_{r+2}, \cdots, a_n = -P_n.$$

Then we may write $f(x)$ as follows:

$$f(x) = x^{n-r}\left[a_0x^r + a_1x^{r-1} + \cdots + a_r - \frac{P_{r+1}}{x} - \frac{P_{r+2}}{x^2} - \cdots - \frac{P_n}{x^{n-r}}\right].$$

The expression $a_0x^r + a_1x^{r-1} + \cdots + a_r$ *increases* as x increases, $(x > 0)$ unless $r = 0$, and then it remains constant; the expression

$$\frac{P_{r+1}}{x} + \frac{P_{r+2}}{x^2} + \cdots + \frac{P_n}{x^{n-r}}$$

diminishes as x increases, $x > 0$. Thus, as x increases from zero onwards, the two expressions cannot be equal more than once. That is, $f(x) = 0$ has only one positive root.

The method of proof will be the same if we suppose the first set of terms negative and the second positive.

Illustration:

$$x^3 + 3x^2 - 4x - 12 = (x + 2)(x + 3)(x - 2) = 0.$$

has only one positive root.

Exercise. In the astronomical problem of three bodies occurs the equation

$$r^5 + (3 - \mu)r^4 + (3 - 2\mu)r^3 - \mu r^2 - 2\mu r - \mu = 0, \qquad 0 < \mu < 1.$$

Show that there is a single positive real root.

5.6 Relations between the roots and the coefficients. In the linear equation and the quadratic equation we have explicit expressions showing the dependence of the roots on the coefficients [see Chapter I]. In this article we will obtain, for equations of higher degree, explicit expressions showing the dependence of the roots on the coefficients.

For simplicity take the coefficient of the highest power of x as plus one. Let the n roots of the equation be $r_1, r_2, \cdots, r_n$. We have the following identity:

$$x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n \equiv (x - r_1)(x - r_2)(x - r_3) \cdots (x - r_n). \tag{1}$$

Let us assume that

$$\begin{aligned}
&(x - r_1)(x - r_2)(x - r_3) \cdots (x - r_n) \\
&\quad \equiv x^n - (r_1 + r_2 + \cdots + r_n)x^{n-1} \\
&\quad + (r_1r_2 + r_1r_3 + \cdots + r_1r_n + r_2r_3 + \cdots + r_{n-1}r_n)x^{n-2} \qquad (2) \\
&\quad - (r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + \cdots + r_{n-2}r_{n-1}r_n)x^{n-3} \\
&\quad + \cdots + (-1)^n(r_1r_2r_3 \cdots r_{n-1}r_n).
\end{aligned}$$

By actual multiplication we have

$$(x - r_1)(x - r_2) \equiv x^2 - (r_1 + r_2)x + r_1r_2$$

$$\begin{aligned}
&(x - r_1)(x - r_2)(x - r_3) \\
&\qquad \equiv x^3 - (r_1 + r_2 + r_3)x^2 + (r_1r_2 + r_1r_3 + r_2r_3)x - r_1r_2r_3,
\end{aligned}$$

so that (2) is true for $n = 2$ and $n = 3$.

If we multiply both sides of (2) by $x - r_{n+1}$, it is easy to see that the resulting equation can be obtained from (2) by replacing n by $n + 1$. Thus by mathematical induction the proof is complete that equation (2) is correct.

Equating coefficients of like powers of x from the left member of (1) and the right member of (2), we have

$$\begin{aligned}
p_1 &= -(r_1 + r_2 + r_3 + \cdots + r_n) \\
p_2 &= r_1r_2 + r_1r_3 + \cdots + r_1r_n + r_2r_3 + \cdots + r_{n-1}r_n \\
p_3 &= -(r_1r_2r_3 + r_1r_2r_4 + r_1r_3r_4 + \cdots + r_{n-2}r_{n-1}r_n) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
p_n &= (-1)^n(r_1r_2r_3 \cdots r_{n-1}r_n).
\end{aligned}$$

These relations between the coefficients and roots may be stated as follows:

Theorem VI: *In a polynomial equation of degree n in x, in which the coefficient of the term of highest degree is plus one,*

(a) *the coefficient of x^{n-1}, with its sign changed, is equal to the sum of the roots;*

(b) *the coefficient of x^{n-2} is the sum of the products of the roots taken two at a time;*

(c) *the coefficient of x^{n-3}, with its sign changed, is equal to the sum of the products of the roots taken three at a time;*

(d) *the coefficient* p_r *of* x^{n-r}, *with its sign changed or not according as* r *is odd or even, is equal to the sum of the products of the roots taken* r *at a time*;

(e) *finally, the product of all of the roots is equal to the constant term or its negative, according as* n *is even or odd.*

Example. Find the cubic equation in p-form whose roots are 2, 3, −4.

$$p_1 = -[2 + 3 + (-4)] = -1;$$

$$p_2 = 2\cdot 3 + 2(-4) + 3(-4) = -14;$$

$$p_3 = -2\cdot 3(-4) = 24.$$

The cubic is therefore

$$x^3 - x^2 - 14x + 24 = 0.$$

Exercises

Construct the equation of lowest degree whose roots are as follows:

1. 2, 3, 4	7. $\pm\sqrt{3}$, 1	13. $\pm\sqrt{2}$, 1, 3
2. 2, 2, −3	8. $\pm\sqrt{3}$, $\pm\sqrt{2}$	14. $\pm\sqrt{3}$, 2, 4
3. 1, 2, 3, 4	9. 1, 2, 3, 4, 5	15. $\pm\sqrt{2}$, $\pm\sqrt{5}$
4. 3, −3, 2	10. 4, 2, −3, −3	16. 3, $1 \pm 2i$
5. 2, −2, 3, −3	11. 4, 5, −3, −2	17. ±2, ±3
6. 2, 2, 2, 1	12. −1, −2, 3, 4	18. −3, −1, 2, 4

5.7 Theorem VII. *Any rational root of an equation*

$$f(x) \equiv x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n = 0 \quad (p_n \neq 0),$$

where $p_1, p_2, \cdots, p_n$ *are integers, is an integer and an exact divisor of* p_n.

If possible let $\frac{a}{b}$ be a root of $f(x) = 0$, where $\frac{a}{b}$ is a fraction in its lowest terms. Then

$$\left(\frac{a}{b}\right)^n + p_1\left(\frac{a}{b}\right)^{n-1} + p_2\left(\frac{a}{b}\right)^{n-2} + \cdots + p_{n-1}\left(\frac{a}{b}\right) + p_n = 0 \quad (3)$$

whence

$$\frac{a^n}{b} = -(p_1a^{n-1} + p_2a^{n-2}b + \cdots + p_{n-1}ab^{n-2} + p_nb^{n-1}). \quad (4)$$

The right-hand member of (4) is an integer, since every term is an integer, while the left-hand member is a fraction in its lowest terms. Thus, the hypothesis that a/b is a root has led to an absurdity.

If r is a rational root, it must then be an integer. We will now prove that it must be a divisor of p_n. If r is an integral root, we have

$$r^n + p_1 r^{n-1} + p_2 r^{n-2} + \cdots + p_{n-1} r + p_n = 0,$$

or

$$-(r^{n-1} + p_1 r^{n-2} + p_2 r^{n-3} + \cdots + p_{n-1}) = \frac{p_n}{r} \tag{5}$$

The left-hand member of (5) is an integer and therefore p_n/r is an integer. This is equivalent to saying that p_n is exactly divisible by r.

Example: Find the roots of $x^4 - 3x^3 + x^2 + 7x - 30 = 0$. The factors of 30 are $\pm1, \pm2, \pm3, \pm5, \pm6, \pm10, \pm15, \pm30$. By trial ±1, $+2$ are not roots. By synthetic division -2 is a root.

$$\begin{array}{rrrrr|l} 1 & -3 & +1 & +7 & -30 & \underline{-2} \\ & -2 & +10 & -22 & +30 & \\ \hline 1 & -5 & +11 & -15 & & \end{array}$$

The quotient is $x^3 - 5x^2 + 11x - 15$. Additional rational roots must be factors of 15; then $\pm2, \pm6, \pm10, \pm30$ are eliminated from further consideration. By trial, -3 is not a root. By synthetic division 3 is a root.

$$\begin{array}{rrrr|l} 1 & -5 & +11 & -15 & \underline{3} \\ & 3 & -6 & +15 & \\ \hline 1 & -2 & +5 & & \end{array}$$

The quotient $x^2 - 2x + 5$ is zero for $x = 1 \pm 2i$. The roots are then $1 \pm 2i, 3, -2$.

5.8 Theorem VIII: *For an algebraic equation all of whose coefficients are integers, an integral root is an exact divisor of the constant term.*

This theorem differs from theorem VII in that in VII $a_0 = 1$, while in this theorem a_0 is not necessarily $= 1$.

If r is a root of the equation

$$a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0 \qquad (a_n \neq 0),$$

then

$$a_0 r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n = 0,$$

whence

$$r(-a_0 r^{n-1} - a_1 r^{n-2} - \cdots - a_{n-1}) = a_n.$$

Since a_0, a_1, $\cdots$, a_n and r are integers, the expression in the parenthesis is an integer and is the quotient obtained by dividing a_n by r. So a_n is exactly divisible by r.

Example. Solve

$$2x^3 \quad 11x^2 + 17x - 6 = 0.$$

The integral divisors of 6 are ± 1, ± 2, ± 3, ± 6. By trial ± 1, -2 are not roots. By synthetic division 2 is a root

$$\begin{array}{rrrr|l} 2 & -\ 11 & +\ 17 & -\ 6 & 2 \\ & 4 & -\ 14 & +\ 6 & \\ \hline 2 & -\ 7 & +\ 3 & +\ 0 & \end{array}$$

To obtain the remaining roots, solve the quadratic equation

$$2x^2 - 7x + 3 = 0.$$

The roots are 2, 3, $\frac{1}{2}$.

Exercises

Find the roots of the following equations:

1. $x^4 - 4x^3 + 4x - 1 = 0$
2. $x^4 - 5x^3 + 11x^2 - 13x + 6 = 0$
3. $x^3 + x^2 + x + 1 = 0$
4. $x^3 - 6x^2 + 11x - 6 = 0$
5. $x^4 - 4x^3 - 8x + 32 = 0$
6. $x^4 - 16x^3 + 86x^2 - 176x + 105 = 0$
7. $x^4 - 9x^3 + 23x^2 - 9x - 18 = 0$
8. $x^4 + 2x^3 - 21x^2 - 22x + 40 = 0$
9. $x^4 + 4x^3 - 34x^2 - 76x + 105 = 0$
10. $x^4 - 8x^3 + 14x^2 + 8x - 15 = 0$
11. $x^5 - 10x^3 + 2x^2 + 17x + 6 = 0$
12. $x^5 - 30x^3 + 34x^2 - 27x - 90 = 0$
13. $x^3 - 28x + 48 = 0$
14. $x^3 + 2x^2 - 23x - 60 = 0$
15. $x^4 - 25x^2 + 60x - 36 = 0$

16. $x^5 - 17x^3 + 12x^2 + 52x - 48 = 0$
17. $x^5 - 41x^3 - 84x^2 + 148x + 336 = 0$
18. $x^6 - 14x^4 + 49x^2 - 36 = 0$
19. $x^6 - 30x^4 + 129x^2 - 100 = 0$
20. $x^8 - 39x^6 + 399x^4 - 1261x^2 + 900 = 0$
21. $6x^3 - 17x^2 - 5x + 6 = 0$
22. $2x^3 + 13x^2 + 17x + 12 = 0$
23. $3x^3 - 40x^2 + 133x - 40 = 0$
24. $3x^3 - 37x^2 + 100x + 84 = 0$
25. $6x^3 - 43x^2 + 71x - 30 = 0$
26. $3x^3 - 35x^2 + 112x - 60 = 0$

5.9 Newton's method of divisors. *Let r be an integral root of*

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-2}x^2 + a_{n-1}x + a_n = 0. \tag{6}$$

Then (6) is exactly divisible by $x - r$. Let the quotient be

$$b_0x^{n-1} + b_1x^{n-2} + b_2x^{n-3} + \cdots + b_{n-2}x + b_{n-1}.$$

The method of synthetic division showed us that

$$\begin{aligned} &a_0 = b_0, \quad a_1 - b_1 = -rb_0, \quad a_2 - b_2 = -rb_1, \cdots \\ &a_{n-1} - b_{n-1} = -rb_{n-2}, \quad a_n = -rb_{n-1}. \end{aligned} \tag{7}$$

By theorem VIII, a_n is exactly divisible by r. This is also seen from the last of equation (7), the quotient being $-b_{n-1}$. The work for the synthetic division was arranged as follows:

$$\begin{array}{l} a_0 + a_1 + a_2 + \cdots + a_{n-2} + a_{n-1} + a_n \,\underline{|\, r} \\ \qquad\; rb_0 + rb_1 + \cdots + rb_{n-3} + rb_{n-2} + rb_{n-1} \\ \hline b_0 + b_1 + b_2 + \cdots + b_{n-2} + b_{n-1} \end{array} \tag{8}$$

Let us rearrange the tabular scheme in (8) in this way:

$$\begin{array}{l} a_0 + a_1 + a_2 + \qquad + a_{n-2} + a_{n-1} + a_n \\ -b_0 - b_1 - b_2 + \qquad - b_{n-2} - b_{n-1} \\ \;\; 0 - rb_0 - rb_1 - \cdots - rb_{n-3} - rb_{n-2} \end{array} \tag{9}$$

In (8) the work progresses from left to right.
In (9) the work progresses from right to left.
In (8) at each step a multiplication is performed.
In (9) at each step a division is performed.

By equations (7) $a_n \div r = -b_{n-1}$. Place $-b_{n-1}$ under a_{n-1}. Then by (7) $a_{n-1} - b_{n-1} = -rb_{n-2}$. Divide $-rb_{n-2}$ by r and place the quotient $-b_{n-2}$ under a_{n-2}. Then by (7) $a_{n-2} - b_{n-2} = -rb_{n-3}$. Continue in this manner. The result of the last division must be $-b_0 = -a_0$. If r is a root, each number that appears below the line in (9) must be exactly divisible by r. If below the line in (9) there turns up a number which is not exactly divisible by r, we know that r is not a root and the computation with respect to r stops at that point.

Example. Find the integral roots of $x^4 - 2x^3 - 13x^2 + 38x - 24 = 0$. The integral divisors of 24 are ± 1, ± 2, ± 3, ± 4, ± 6, ± 8, ± 12, ± 24.

Try $x = 24$:

$$\begin{array}{rrrrr|l} 1 & -\ 2 & -\ 13 & +\ 38 & -\ 24 & \underline{24} \\ & & & -\ 1 & & \\ \hline & & & +\ 37 & & \end{array}$$

24 is rejected, since 37 is not divisible by 24.

Try $x = 12$:

$$\begin{array}{rrrrr|l} 1 & -\ 2 & -\ 13 & +\ 38 & -\ 24 & \underline{12} \\ & & +\ 3 & -\ 2 & & \\ \hline & & -10 & +\ 36 & & \end{array}$$

12 is rejected, since -10 is not divisible by 12.

Try $x = 6$:

$$\begin{array}{rrrrr|l} 1 & -\ 2 & -\ 13 & +\ 38 & -\ 24 & \underline{6} \\ & & & -\ 4 & & \\ \hline & & & +\ 34 & & \end{array}$$

6 is rejected, since 34 is not divisible by 6.

Try $x = 4$:

$$\begin{array}{rrrrr|l} 1 & -\ 2 & -\ 13 & +\ 38 & -\ 24 & 4 \\ & & +\ 8 & -\ 6 & & \\ \hline & & -\ 5 & +\ 32 & & \end{array}$$

4 is rejected, since -5 is not divisible by 4.

Try $x = 3$:

$$\begin{array}{rrrrr|l} 1 & -\ 2 & -\ 13 & +\ 38 & -\ 24 & 3 \\ -1 & -\ 1 & +\ 10 & -\ 8 & & \\ \hline 0 & -\ 3 & -\ 3 & +\ 30 & & \end{array}$$

3 is a root. We now have to solve the cubic

$$x^3 + x^2 - 10x + 8 = 0.$$

The factors of 8 are ± 1, ± 2, ± 4, ± 8. $+4$ has already been tried and rejected.

Try $x = 8$:

$$\begin{array}{rrrr|l} 1 & +1 & -10 & +8 & \underline{8} \\ & & 1 & & \\ \hline & & -9 & & \end{array}$$

8 is rejected, since -9 is not divisible by 8.

Try $x = 2$:

$$\begin{array}{rrrr|l} 1 & +1 & -10 & +8 & \underline{2} \\ -1 & -3 & +4 & & \\ \hline 0 & -2 & -6 & & \end{array}$$

$+2$ is a root. We now have to solve $x^2 + 3x - 4 = 0$.

The roots of this quadratic are -4, $+1$.

The roots of the given quartic are then -4, 1, 2, 3.

Exercises

Find the integral roots of the following equations.

1. $x^4 + x^3 - 2x^2 + 4x - 24 = 0$
2. $x^4 - 2x^3 - 19x^2 + 68x - 60 = 0$
3. $x^3 - 15x^2 + 74x - 120 = 0$
4. $x^3 - 10x^2 + x + 120 = 0$
5. $x^4 - 11x^3 + 6x^2 + 184x - 320 = 0$
6. $x^4 - 99x^2 + 130x + 1200 = 0$
7. $x^4 - 85x^2 + 324 = 0$
8. $x^4 + x^3 - 63x^2 - 64x - 64 = 0$
9. $x^4 - 15x^3 + 63x^2 - 62x + 48 = 0$
10. $x^4 - 5x^3 + 16x^2 - 50x + 60 = 0$
11. $x^4 - 143x^2 - 144 = 0$
12. $x^4 - 32x^2 - 144 = 0$
13. $x^4 - 106x^2 + 600 = 0$
14. $x^4 - 73x^2 + 576 = 0$
15. $x^4 - 27x^2 + 14x + 120 = 0$
16. $x^4 - 9x^3 + 148x - 240 = 0$
17. $5x^4 - 43x^3 + 684x - 1296 = 0$
18. $x^5 + 5x^4 - 13x^3 - 65x^2 + 36x + 180 = 0$
19. $x^5 - 5x^4 - 37x^3 + 185x^2 + 36x - 180 = 0$
20. $x^5 - 3x^4 - 39x^3 + 104x^2 + 108x + 144 = 0$

5.10 Method of limiting the number of divisors. If r is a root of $f(x) = 0$, then $f(x)$ is exactly divisible by $x - r$, that is

$$f(x) \equiv (x - r)Q(x).$$

This equation is true for every value of x. Give to x the integral value a, then

$$f(a) = (a - r)Q(a).$$

In other words, *if r is a root of $f(x) = 0$, where $f(x)$ has integral coefficients, then $f(a)$ is exactly divisible by $a - r$ or $r - a$.*
In particular, if $a = +1$ or -1, we have

$$f(1) = (1 - r)Q(1)$$

$$f(-1) = (-1 - r)Q(-1).$$

It is easy to compute $f(1)$ and $f(-1)$. Do this. Then *if r is a root of $f(x) = 0$, $f(1)$ is exactly divisible by $r - 1$. Also if r is a root of $f(x) = 0$, $f(-1)$ is exactly divisible by $r + 1$.*

Example: Find the integral roots of

$$x^5 - 23x^4 + 160x^3 - 281x^2 - 257x - 440 = 0.$$

The factors of 440 are ±1, ±2, ±4, ±5, ±8, ±10, ±11, ±20, ±22, ±40, ±44, ±55, ±88, ±110, ±220, ±440.
We have

$$f(1) = -840, \quad \text{and} \quad f(-1) = -648.$$

We exclude all of the above divisors, which, when diminished by 1, do not divide 840. This leaves out ±440, ±220, ±110, ±88, ±55, ±44, ±40, -22, $+20$, ±10, -8.

We omit all of the above divisors, which, when increased by 1, do not divide 648. This gives us as additional exclusions $+22$, -20, -11, $+4$, and we have left only ±1, ±2, ±5, 8, 11, -4.

We find that 5, 8, 11 are roots and the resulting quotient is $x^2 + x + 1$. Hence the given equation is equivalent to

$$(x - 5)(x - 8)(x - 11)(x^2 + x + 1) = 0,$$

and the only integral roots are 5, 8, 11.

Exercises

Find the integral roots of the following equations.

1. $x^5 - 29x^4 - 31x^3 + 31x^2 - 32x + 60 = 0$
2. $x^5 - 2x^4 - 15x^3 - 12x^2 + 44x + 80 = 0$

3. $x^4 - 11x^3 + 16x^2 + 54x + 80 = 0$
4. $x^4 + 2x^3 - 27x^2 - 28x - 60 = 0$
5. $x^5 - 14x^4 + 57x^3 - 96x^2 + 212x - 160 = 0$
6. $x^5 - 18x^4 + 89x^3 - 106x^2 + 174x - 140 = 0$
7. $x^4 - 11x^3 - 28x^2 + 260x + 528 = 0$
8. $x^4 - 61x^2 + 900 = 0$
9. $x^6 - 27x^4 + 42x^2 + 200 = 0$
10. $x^6 + 7x^5 - 22x^4 - 238x^3 - 183x^2 + 1575x + 2700 = 0$

5.11 Theorem IX. *If an algebraic equation, all of whose coefficients are integers,*

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \quad (a_0 > 0,\ a_n \neq 0), \tag{10}$$

has a rational root p/q, where p and q are integers which have no common divisor other than unity, then p is an exact divisor of a_n, and q is an exact divisor of a_0.

In (10) put $x = p/q$ and multiply by q^n. We find

$$a_0p^n + a_1p^{n-1}q + \cdots + a_{n-1}pq^{n-1} + a_nq^n = 0,$$

which may be written either

$$p(a_0p^{n-1} + a_1p^{n-2}q + \cdots + a_{n-1}q^{n-1}) = -a_nq^n, \tag{11}$$

or

$$q(a_1p^{n-1} + \qquad + a_{n-1}pq^{n-2} + a_nq^{n-1}) = -a_0p^n. \tag{12}$$

Equation (11) shows plainly that p is a factor of the left-hand member; hence p must be a factor of the right-hand member. But by hypothesis p and q have no common factor other than unity. Then p and q^n can have no common factor other than unity. But p must be a factor of the right-hand side. Therefore p must be a factor of a_n. In like manner from (12) we can prove that q must be a factor of a_0.

If one multiplies (10) by a_0^{n-1} and then makes the substitution $y = a_0x$, (10) will be changed to an equation in y in which the coefficient of y^n is plus one. If $a_n = 1$, try the substitution $y = 1/x$.

Exercises

Find all of the rational roots of the following equations.

1. $6x^3 + 7x^2 - 9x + 2 = 0$
2. $4x^4 - 13x^2 + 9 = 0$

3. $6x^3 - x^2 - 19x - 6 = 0$
4. $4x^3 - 17x^2 + 9x + 18 = 0$
5. $3x^3 + 7x^2 - 7x - 3 = 0$
6. $4x^4 - 20x^3 + 23x^2 + 5x - 6 = 0$
7. $2x^4 - 9x^3 + 6x^2 + 11x - 6 = 0$
8. $12x^3 + 4x^2 - 53x + 30 = 0$
9. $6x^4 + 5x^3 - 29x^2 - 20x + 20 = 0$
10. $12x^3 - 4x^2 - 3x + 1 = 0$
11. $36x^5 - 36x^4 - 13x^3 + 13x^2 + x - 1 = 0$

5.12 Determination of multiple roots. If $f(x) \equiv (x - r)^m \phi(x)$ $[\phi(r) \neq 0]$, then r is said to be a root of $f(x)$ of multiplicity (order) m: that is, by definition, r occurs as a root m times.

If $r_1, r_2, \cdots, r_n$ are the roots of $f(x) = 0$, then

$$f(x) \equiv (x - r_1)(x - r_2) \cdots (x - r_n),$$

and

$$f(h + x) \equiv (h + x - r_1)(h + x - r_2) \cdots (h + x - r_n).$$

By Taylor's theorem, §4.3, we have

$$f(h + x) = f(x) + f'(x)h + \frac{f''(x)}{1.2} h^2 + \cdots + h^n. \tag{13}$$

The coefficient of h in these two expressions for $f(h + x)$ must be identical, and so

$$\begin{aligned} f'(x) = (x - r_2)(x - r_3) \cdots (x - r_n) \\ + (x - r_1)(x - r_3) \cdots (x - r_n) + \cdots \\ + (x - r_1)(x - r_2) \cdots (x - r_{n-1}), \end{aligned}$$

where the right-hand member is the sum of n terms, each the product of $n - 1$ linear factors $x - r_i$.

This value of $f'(x)$, for x not a root, may be written

$$f'(x) = \frac{f(x)}{x - r_1} + \frac{f(x)}{x - r_2} + \cdots + \frac{f(x)}{x - r_n}. \tag{14}$$

If the factor $(x - r_1)^m$ occurs in $f(x)$, *i.e.*, if $r_1 = r_2 = \cdots = r_m$, we have

$$f'(x) = \frac{mf(x)}{x - r_1} + \frac{f(x)}{x - r_{m+1}} + \cdots + \frac{f(x)}{x - r_n}$$

Each term in the right-hand member of this equation will have $(x - r_1)^m$ as a factor, except the first, which will have $(x - r_1)^{m-1}$ as a factor; hence $(x - r_1)^{m-1}$, but not $(x - r_1)^m$, is a factor of $f'(x)$. Thus we have proved the following:

Theorem X: *A multiple root of order m of $f(x) = 0$ is a multiple root of order $m - 1$ of the first derived equation $f'(x) = 0$, and so $f(x)$ and $f'(x)$ have the common factor $(x - r)^{m-1}$.*

Corollary I. *Any root which occurs m times in $f(x) = 0$ occurs in degrees of multiplicity diminishing by unity in the first $m - 1$ derived functions.*

Corollary II. *r is a root of order m of $f(x) = 0$, if and only if $f(r) = f'(r) = \cdots = f^{(m-1)}(r) = 0$, but $f^{(m)}(r) \neq 0$.*

In (13) replace x by r and h by $x - r$; then (13) becomes

$$f(x) = f(r) + f'(r)(x - r) + \cdots \\ + \frac{f^{(m-1)}(r)}{\underline{|m - 1}} (x - r)^{m-1} + \frac{f^{(m)}(r)}{\underline{|m}} (x - r)^m + \cdots + (x - r)^n. \tag{15}$$

From this form of $f(x)$ it is clear that, if

$$f(r) = f'(r) = \cdots = f^{(m-1)}(r) = 0, \text{ but } f^{(m)}(r) \neq 0,$$

$(x - r)^m$ is a factor of $f(x)$ and r is a root of multiplicity m.

If $f(r) = f'(r) = \cdots = f^{(m-s)}(r) = 0$ but $f^{(m-s+1)}(r) \neq 0$, $s > 1$, then $(x - r)^{m-s+1}$ is a factor of $f(x)$, but $(x - r)^m$ is not a factor and so the root r does not have the multiplicity m.

Let us now suppose that $(x - r)^{m-1}$ is a common factor of $f(x)$ and $f'(x)$ but that $(x - r)^m$ is not a common factor, $m > 1$. In (15) replace $f(x)$ by $f'(x)$ and $f^{(s)}(x)$ by $f^{(s+1)}(x)$; then

$$f'(x) = f'(r) + f''(r)(x - r) + \cdots \\ + \frac{f^{(m)}(r)}{\underline{|m - 1}} (x - r)^{m-1} + \frac{f^{(m+1)}(r)}{\underline{|m}} (x - r)^m + \cdots. \tag{16}$$

Since $(x - r)^{m-1}$ is a factor of $f'(x)$, from (16) we have $f'(r) = f''(r) = \cdots = f^{(m-1)}(r) = 0$. Since $x - r$ is a factor of $f(x)$, we must have $f(r) = 0$. From (15) $f(x)$ has the factor $(x - r)^m$, which by hypothesis is not a factor of $f'(x)$. It follows that in (16) $f^{(m)}(r) \neq 0$. Hence r is a root of $f(x) = 0$ of order m.

This completes the proof of the following theorem:

Theorem XI: *If $f(x)$ and $f'(x)$ have a H.C.F., which is not a mere constant, this H.C.F. will be a polynomial [which may be designated by $d(x)$]. Any root of $f(x) = 0$ of order $m > 1$ will be a root of $f'(x) = 0$ of order $m - 1$ and hence a root of $d(x) = 0$ of order $m - 1$; conversely, any root of $d(x) = 0$ of order $m - 1$ is a root of $f'(x) = 0$ of order $m - 1$ and hence a root of $f(x) = 0$ of order m.*

In order, therefore, to find whether any proposed equation has equal roots and to determine such roots when they exist, we determine the H.C.F. of $f(x)$ and $f'(x)$. If this H.C.F. is not a mere constant, let it be designated by $d(x)$. Determine the roots of $d(x) = 0$ and the multiplicity of each.

Example. Find the multiple roots of $x^3 - x^2 - 8x + 12 = 0$. The H.C.F. of $f(x) = x^3 - x^2 - 8x + 12$ and $f'(x) = 3x^2 - 2x - 8$ is $x - 2$. Hence $(x - 2)^2$ is a factor of $f(x)$. The other factor is $x + 3$, and the roots of the given equation are 2, 2, -3.

Exercises

Find the multiple roots of the following equations:

1. $x^4 - 2x^3 - 3x^2 + 8x - 4 = 0$
2. $x^3 + x^2 - x - 1 = 0$
3. $x^4 - x^3 - x + 1 = 0$
4. $x^3 + 5x^2 + 8x + 4 = 0$
5. $x^4 + 2x^3 + 5x^2 + 8x + 4 = 0$
6. $x^4 + 2x^3 - 2x - 1 = 0$
7. $x^4 + 5x^3 + 6x^2 - 4x - 8 = 0$
8. $x^4 - 9x^2 - 4x + 12 = 0$
9. $x^4 - 9x^2 + 4x + 12 = 0$
10. $x^4 - 2x^3 - 11x^2 + 12x + 36 = 0$
11. $x^4 + 3x^3 - 7x^2 - 15x + 18 = 0$
12. $x^4 - 2x^3 - 12x^2 + 18x + 27 = 0$
13. $x^6 + x^5 - x^4 - 2x^3 - x^2 + x + 1 = 0$
14. $x^5 - x^4 + 2x^3 - 2x^2 + x - 1 = 0$
15. $x^5 - 2x^4 - 6x^3 + 8x^2 + 9x + 2 = 0$
16. $x^6 + x^5 - 5x^4 - 6x^3 + 3x^2 + 9x + 9 = 0$

17. Show that the equation $x^n - a^n = 0$ cannot have equal roots.

18. Determine the condition that the cubic $x^3 + 3Hx + G = 0$ should have two equal roots.

5.13 Rolle's theorem* for polynomials. *Between two consecutive real roots a and b of* $f(x) = 0$ *there is at least one real root of* $f'(x) = 0$.

Let the curve in the figure be the graph of $y = f(x)$. The real roots of $f(x) = 0$ are the x coordinates of the points where the curve crosses the x-axis, namely the points P, Q, R, S (Fig. 19). At P there is a double root of $f(x) = 0$. Both $f(x)$ and $f'(x)$ are polynomials and hence are continuous functions of the variable x. The continuity of $f'(x)$ assures us that the tangent line to the curve turns continuously about the point of tangency as the point of tangency travels along the curve.

As x increases from a to b, $f(x)$, varying continuously from $f(a) = 0$ to $f(b) = 0$, must either begin by decreasing and then

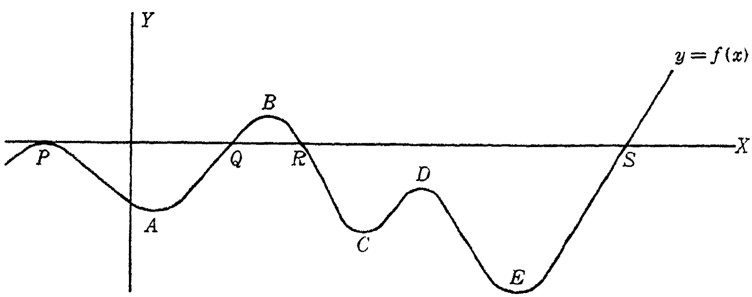
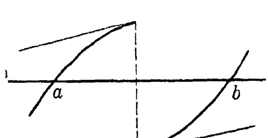


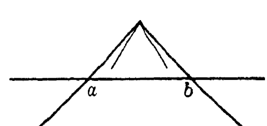
Fig. 19

increasing, as from P to Q, or must begin by increasing and then decreasing, as from Q to R; or $f(x)$ may decrease and then increase (increase and then decrease) several times between a and b, as in Fig. 19 between R and S. Then $f(x)$ must have at least one

* For the general case of real variables, this theorem is stated as follows: *If* $f(x)$ *and* $f'(x)$ *are continuous in an interval* (a, b) *and* $f(a) = f(b) = 0$, *then* $f'(x_1) = 0$ *for some* x_1, *where* $a < x_1 < b$. The necessity for these conditions is illustrated by the following diagrams.



$f(x)$ not continuous



$f(x)$ continuous
$f'(x)$ not continuous

maximum or minimum between a and b. In the figure there is a minimum at A between P and Q; a maximum at B between Q and R; and a minimum at C, a maximum at D, a minimum at E, between R and S. The value α of x for which a maximum (minimum) of $f(x)$ occurs is a real root of $f'(x) = 0$, since the slope of the tangent line at a maximum or minimum is zero.

In the figure there is one real root of $f'(x) = 0$ between P and Q, and one real root between Q and R. There are three real roots of $f'(x) = 0$ between R and S; these roots are the x-coordinates of the points C, D, E. From the figure it appears that the number of real roots of $f'(x) = 0$ between any two consecutive real roots of $f(x) = 0$ is always odd, multiple roots being counted with the proper multiplicity.

Corollary I: *Between two consecutive real roots α and β of $f'(x) = 0$, there cannot be more than one real root of $f(x) = 0$.*

For, if $f(x) = 0$ had two real roots a and b, between α and β, then by Rolle's theorem $f'(x) = 0$ would have a real root γ between a and b. But then $f'(x) = 0$ would have a real root γ between α and β which is contrary to the hypothesis that α and β are consecutive roots.

Corollary II: *Not more than one real root of $f(x) = 0$ is greater than the greatest real root of $f'(x) = 0$; and not more than one real root of $f(x) = 0$ is less than the smallest real root of $f'(x) = 0$.*

For, if $f(x) = 0$ had two real roots a and b greater (less) than the greatest (smallest*) real root α of $f'(x) = 0$, then by Rolle's theorem $f'(x) = 0$ would have a real root $\gamma > \alpha (< \alpha)$ between a and b, which is contrary to the hypothesis that α is the greatest (smallest) real root of $f'(x) = 0$.

Example. Locate the roots of $x^3 - 6x^2 + 9x - 2 = 0$.

$$f(x) = x^3 - 6x^2 + 9x - 2$$

$$\tfrac{1}{3}f'(x) = x^2 - 4x + 3.$$

The roots of $f'(x) = 0$ are 1 and 3.

$$f(-\infty) < 0; \qquad f(1) = 2; \qquad f(3) = -2; \qquad f(+\infty) > 0.$$

Hence there is one root less than 1, one root greater than 3, and one root between 1 and 3. Since $f(0) = -2$ and $f(1) = 2$, the

* The reader must keep in mind that -5 is smaller than -1.

root less than 1 is between 0 and 1. Since $f(4) = 2$ and $f(3) = -2$, the root greater than 3 is between 3 and 4.

Exercises

Locate the roots of

1. $2x^3 - 15x^2 + 36x - 27 = 0$
2. $2x^3 - 21x^2 + 60x - 40 = 0$
3. $x^3 - 12x^2 + 36x - 16 = 0$
4. $3x^4 - 8x^3 - 6x^2 + 24x - 10 = 0$
5. $3x^4 - 4x^3 - 24x^2 + 48x - 20 = 0$
6. $3x^5 - 50x^3 + 135x + 10 = 0$
7. $3x^5 - 25x^3 + 60x + 25 = 0$
8. $x^4 - 56x^2 + 192x - 150 = 0$
9. Show that $x^3 + x^2 - 10x + 9 = 0$ has two roots between 1 and 2.

Chapter VI

TRANSFORMATION OF EQUATIONS

6.1 Introduction. We have learned something about how to find the rational roots of an equation. However, we have not considered the matter of finding the real irrational roots.

Sometimes the solution of an equation is facilitated by transforming it into another equation whose roots are related to the roots of the given equation in some specified manner. The transformations given in this chapter will assist us in finding rational roots and real irrational roots. We shall make use of the following transformation of

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 > 0).$$

If in $f(x) = 0$ we put $x = -y$, we obtain an equation in y whose roots are those of $f(x) = 0$ with their signs changed. This transformation enables us to find the negative roots of a given equation by finding the positive roots of the transformed equation.

If in $f(x) = 0$ we put $x = y/m$, we obtain an equation in y whose roots are those of $f(x) = 0$, each multiplied by m. This transformation enables us to find the rational roots of a given equation by finding first the integral roots of the transformed equation.

If in $f(x) = 0$ we put $x = y + h$, we obtain an equation in y whose roots are those of $f(x) = 0$ each diminished by h. This transformation enables us to find for a given equation a root between h and $h + 1$ by first finding the corresponding root of the transformed equation that is between 0 and 1. The labor of computing the roots is much reduced as will be seen when we come to the solution by Horner's method, given in a later chapter.

If in $f(x) = 0$, with $f(0) \neq 0$, we put $x = \frac{1}{y}$, we obtain an equation in y whose roots are the reciprocals of those of $f(x) = 0$. This is of use in finding rational roots of an equation in which the constant term is unity and the coefficient of the term of highest degree is an integer greater than one.

The transformation that is applied to a reciprocal equation of even degree enables us to replace the given equation by another whose degree is half that of the given equation. The solutions of the transformed equation together with that of a quadratic furnish the solutions of the given equation.

6.2 Transformations. 1. *To transform an equation into another whose roots are those of the given equation with their signs changed:*

Let the roots of the given equation be $r_1, r_2, \cdots, r_n$. Then

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$
$$\equiv a_0(x - r_1)(x - r_2) \cdots (x - r_n).$$

Putting $x = -y$, we have

$$a_0(-y)^n + a_1(-y)^{n-1} + \cdots + a_{n-1}(-y) + a_n$$
$$\equiv a_0(-y - r_1)(-y - r_2) \cdots (-y - r_n).$$

Whether n is odd or even, this simplifies into

$$a_0y^n - a_1y^{n-1} + a_2y^{n-2} - \cdots + (-1)^n a_n$$
$$\equiv a_0(y + r_1)(y + r_2) \cdots (y + r_n) = 0.$$

The roots of this last equation are $-r_1, -r_2, \cdots, -r_n$.

We have, therefore, the following rule:

Rule.—In order to form an equation whose roots are the roots of a given equation with their signs changed, change the sign of the coefficient of every other term beginning with the coefficient of x^{n-1}. If any power of x is missing it is here regarded as supplied with a zero coefficient.

2. *To transform an equation into another whose roots are those of the given equation, each multiplied by a given constant m:*

Let the roots of the given equation be $r_1, r_2, \cdots, r_n$. Then

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n$$
$$\equiv a_0(x - r_1)(x - r_2) \cdots (x - r_n).$$

Put $x = y/m$. We have

$$a_0\left(\frac{y}{m}\right)^n + a_1\left(\frac{y}{m}\right)^{n-1} + \cdots + a_{n-1}\left(\frac{y}{m}\right) + a_n$$
$$\equiv \left(\frac{y}{m} - r_1\right)\left(\frac{y}{m} - r_2\right) \cdots \left(\frac{y}{m} - r_n\right).$$

Multiply both sides of the last equation by m^n. We have

$$a_0y^n + a_1my^{n-1} + a_2m^2y^{n-2} + \cdots + a_{n-1}m^{n-1}y + a_nm^n$$
$$\equiv a_0(y - mr_1)(y - mr_2) \cdots (y - mr_n) = 0.$$

The roots of this last equation are $mr_1, mr_2, \cdots, mr_n$.

We have, therefore, the following rule:

Rule.—In order to form an equation whose roots are m times the roots of a given equation, multiply the successive coefficients beginning with the coefficient of x^{n-1} *by* $m, m^2, \cdots, m^n$ *respectively. If any power of x is missing, it is here regarded as supplied with a zero coefficient.*

Exercises

Form equations whose roots are the roots of the following equations with their signs changed.

1. $x^3 - 5x^2 - 4x + 3 = 0$
2. $x^3 + 2x - 7 = 0$
3. $x^3 + 1 = 0$
4. $-4x^3 + 2x^2 - 3x - 5 = 0$
5. $2x^4 + 3x^2 - 7 = 0$
6. $3x^4 - 4x^3 + 2x - 5 = 0$
7. $x^5 - x^4 + 2x^3 - 3x^2 + 5x - 4 = 0$
8. $5x^6 - 3x^3 + x^2 - 7 = 0$
9. $7x^4 - 4x^3 + 5x^2 + 3x - 2 = 0$
10. $x^5 - x^2 + x - 3 = 0$
11. $3x^6 - 5x^3 + 7x - 5 = 0$
12. $2x^5 - 3x^4 + 5x^2 - 6 = 0$

Obtain equations whose roots are equal to the roots of the following equations multiplied by the number placed opposite.

13. $5x^3 - x^2 + 3x - 1 = 0$ (2)
14. $2x^4 - 3x^3 + 5x - 6 = 0$ (2)
15. $3x^4 - x^3 + x^2 - 2x + 1 = 0$ (3)
16. $x^4 - 2x^2 + 5 = 0$ (−2)
17. $x^3 + x^2 + x + 1 = 0$ (4)
18. $3x^4 + 5x^3 + 7x - 2 = 0$ (−1)
19. $2x^3 - 3x^2 - x + 1 = 0$ (4)
20. $5x^4 + 3x^2 + x + 2 = 0$ (5)
21. $x^6 + x + 1 = 0$ (2)
22. $3x^5 + 5x^2 - 1 = 0$ (3)

Obtain equations whose roots are the roots of the following equations multiplied by the smallest positive value of m which will make the coefficients of the resulting equation integers when the coefficient of the highest power of x is $+1$.

23. $5x^3 + 3x^2 + 2x - 4 = 0$
24. $x^3 + \frac{1}{2}x^2 - \frac{1}{3}x + \frac{1}{4} = 0$
25. $3x^4 - 5x^3 + 2x - 1 = 0$
26. $x^3 + \frac{1}{3}x^2 - \frac{1}{4}x + \frac{2}{9} = 0$
27. $x^4 - \frac{1}{5}x^2 + \frac{1}{10}x - \frac{1}{8} = 0$
28. $x^4 - \frac{1}{5}x^3 + \frac{2}{25}x^2 - \frac{7}{125} = 0$
29. $4x^4 + 3x^3 + 2x^2 + x + 1 = 0$
30. $2x^4 - 5x^2 + 3x - 4 = 0$
31. $x^5 - \frac{1}{2}x^4 + \frac{1}{3}x^3 + \frac{1}{4}x^2 - \frac{4}{27} = 0$
32. $x^6 - \frac{1}{5}x^3 + \frac{2}{25}x - \frac{3}{125} = 0$
33. $x^4 - \frac{1}{6}x^3 - \frac{1}{60}x^2 + \frac{1}{200}x + \frac{1}{2250} = 0$

6.3 To transform an equation into another whose roots are the reciprocals of the roots of the given equation:

Let the roots of the given equation be $r_1, r_2, \cdots, r_n$. Then

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$
$$\equiv a_0(x - r_1)(x - r_2) \cdots (x - r_n). \quad (a_0, a_n \neq 0)$$

Put $x = \dfrac{1}{y}$. We have

$$\frac{a_0}{y^n} + \frac{a_1}{y^{n-1}} + \cdots + \frac{a_{n-1}}{y} + a_n$$
$$\equiv a_0\left(\frac{1}{y} - r_1\right)\left(\frac{1}{y} - r_2\right) \cdots \left(\frac{1}{y} - r_n\right)$$
$$= \frac{a_0}{y^n}(1 - r_1y)(1 - r_2y) \cdots (1 - r_ny)$$
$$= \frac{a_0r_1r_2 \cdots r_n}{y^n}\left(\frac{1}{r_1} - y\right)\left(\frac{1}{r_2} - y\right) \cdots \left(\frac{1}{r_n} - y\right)$$
$$= \frac{a_n}{y^n}\left(y - \frac{1}{r_1}\right)\left(y - \frac{1}{r_2}\right) \cdots \left(y - \frac{1}{r_n}\right)$$

since $$\frac{a_n}{a_0} = (-1)^n r_1r_2 \cdots r_n.$$

Whence, multiplying by y^n, we have

$$a_ny^n + a_{n-1}y^{n-1} + \cdots + a_1y + a_0$$
$$\equiv a_n\left(y - \frac{1}{r_1}\right)\left(y - \frac{1}{r_2}\right) \cdots \left(y - \frac{1}{r_n}\right).$$

Thus we see that if in the given equation we replace x by $\frac{1}{y}$ and multiply by y^n, the resulting polynomial in y equated to zero will have for roots the reciprocals of the roots of the given equation.

Exercises

Obtain the equations whose roots are the reciprocals of the roots of the following equations.

1. $x^2 + 2x + 1 = 0$
2. $x^3 + 1 = 0$
3. $3x^3 + 5x^2 - 7x + 4 = 0$
4. $2x^4 - 3x^2 - 5x + 7 = 0$
5. $2x^4 - 5x^3 - 5x + 2 = 0$
6. $2x^5 - 3x^4 + 5x^3 + 5x^2 - 3x + 2 = 0$
7. $5x^4 - 7x^3 + 7x - 5 = 0$
8. $x^5 + x^4 + 3x^2 + x + 1 = 0$
9. $x^3 + 3x^2 - 7x + 5 = 0$
10. $3x^4 + 5x^3 - 5x - 3 = 0$

6.4 To transform an equation into another whose roots are those of the given equation diminished (increased) by a constant h. In the given equation

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 \neq 0), \quad (1)$$

put $x = y + h$. The resulting equation in y will have roots less or greater by $|h|$ than the roots of the given equation according as h is positive or negative. We now have

$$a_0(y + h)^n + a_1(y + h)^{n-1} + \cdots + a_{n-1}(y + h) + a_n = 0.$$

Expanding the binomial powers and arranging in powers of y, we represent the transformed equation in y by

$$A_0y^n + A_1y^{n-1} + \cdots + A_{n-1}y + A_n = 0 \qquad (A_0 = a_0 \neq 0).$$

Since $y = x - h$, this is equivalent to

$$A_0(x - h)^n + A_1(x - h)^{n-1} + \cdots + A_{n-1}(x - h) + A_n = 0,$$

which must be identical with the given equation (1). Now

$$A_0(x - h)^n + A_1(x - h)^{n-1} + \cdots + A_{n-1}(x - h) + A_n$$
$$\equiv [A_0(x - h)^{n-1} + A_1(x - h)^{n-2} + \cdots + A_{n-1}]\,(x - h) + A_n.$$

This shows that if we divide the left member of (1) by $x - h$, the remainder is A_n. If we divide the quotient by $x - h$, the remainder is A_{n-1}. Dividing this second quotient by $x - h$, the remainder is A_{n-2}. Continuing in this way $A_0 = a_0$ is the last quotient and A_1 the last remainder. We have, therefore, the following rule.

Rule.—In order to form the equation whose roots are the roots of a given equation each diminished by h, divide the given equation by $x - h$ and denote the remainder by A_n. Divide the quotient by $x - h$, and denote the remainder by A_{n-1}. Continue this process to n divisions. The last quotient, A_0, and the remainders, $A_1, A_2, \cdots, A_n$, are the coefficients of the transformed equation which is then

$$A_0y^n + A_1y^{n-1} + A_2y^{n-2} + \cdots + A_{n-1}y + A_n = 0 \qquad (A_0 = a_0).$$

Example 1. Form an equation whose roots are the roots of

$$3x^4 - 4x^3 - 5x^2 + x + 7 = 0$$

each diminished by 2. Use synthetic division. Divide by $x - 2$.

$$\begin{array}{rrrrr|lcl}
3 & -\ 4 & -\ 5 & +\ 1 & +\ 7 & \underline{2} & & \\
 & 6 & +\ 4 & -\ 2 & -\ 2 & & & \\
\hline
3 & +\ 2 & -\ 1 & -\ 1 & +\ 5 & & A_4 = & 5 \\
 & +\ 6 & +\ 16 & +\ 30 & & & & \\
\hline
3 & +\ 8 & +\ 15 & +\ 29 & & & A_3 = & 29 \\
 & +\ 6 & +\ 28 & & & & & \\
\hline
3 & +\ 14 & +\ 43 & & & & A_2 = & 43 \\
 & +\ 6 & & & & & & \\
\hline
3 & +\ 20 & & & & & A_1 = & 20
\end{array}$$

$3y^4 + 20y^3 + 43y^2 + 29y + 5 = 0$ is the required equation.

Example 2. Form an equation whose roots are the roots of $3x^4 - 4x^3 - 5x^2 + x + 7 = 0$ each increased by 2. To increase a root by 2 is the same as diminishing the roots by -2. Divide by $x + 2$.

$$\begin{array}{rrrrr|lcr}
3 & -\ 4 & -\ 5 & +\ 1 & +\ 7 & \underline{-2} & & \\
 & -\ 6 & +\ 20 & -\ 30 & +\ 58 & & & \\
\hline
3 & -\ 10 & +\ 15 & -\ 29 & +\ 65 & & A_4 = & 65 \\
 & -\ 6 & +\ 32 & -\ 94 & & & & \\
\hline
3 & -\ 16 & +\ 47 & -\ 123 & & & A_3 = & -123 \\
 & -\ 6 & +\ 44 & & & & & \\
\hline
3 & -\ 22 & +\ 91 & & & & A_2 = & 91 \\
 & -\ 6 & & & & & & \\
\hline
3 & -\ 28 & & & & & A_1 = & -28
\end{array}$$

$3y^4 - 28y^3 + 91y^2 - 123y + 65 = 0$ is the required equation.

Exercises

Obtain equations whose roots are the roots of the following equations diminished by the number opposite.

1. $x^4 - 5x^3 + 7x^2 - 17x + 11 = 0$ (4)
2. $x^5 + 4x^3 - x^2 + 11 = 0$ (3)
3. $4x^5 - 2x^3 + 7x - 3 = 0$ (−2)
4. $3x^4 + 7x^3 - 15x^2 + x - 2 = 0$ (−7)
5. $2x^4 + 3x^3 + 4x^2 - 5x - 4 = 0$ (1)
6. $x^4 - 2x^3 - 3x^2 + 8x - 4 = 0$ (1)
7. $x^4 - 8x^3 + 14x^2 - 3x + 2 = 0$ (2)
8. $x^4 - 12x^3 + 25x^2 + 7x + 2 = 0$ (3)
9. $x^3 - 15x^2 + 48x + 8 = 0$ (5)
10. $x^3 - 9x^2 + 15x + 6 = 0$ (3)
11. $x^4 - 5x^3 + 6x^2 - 3x + 8 = 0$ (2)

6.5 Removal of the term a_1x^{n-1}. The transformation $x = y - \dfrac{a_1}{na_0}$ will remove the term of degree $n - 1$. For, let us put $x = y + b$ in

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 > 0).$$

The transformed equation in y is

$$a_0y^n + (na_0b + a_1)y^{n-1} + \cdots = 0.$$

The term containing y^{n-1} will be absent if $b = -\dfrac{a_1}{na_0}$.

Example. Let us remove the second degree term from $x^3 + 6x^2 + x - 2 = 0$. Putting $x = y - \frac{6}{3} = y - 2$, we have

$$(y - 2)^3 + 6(y - 2)^2 + (y - 2) - 2 \equiv y^3 - 11y + 12 = 0.$$

Exercises

Remove the term which is of degree less by one than that of the term of highest degree.

1. $x^3 + 12x^2 + x - 100 = 0$
2. $x^4 + 4x^3 - 8x - 4 = 0$
3. $x^3 - 6x^2 + 7 = 0$
4. $x^4 - 8x^3 + 64x - 70 = 0$
5. $x^3 - 3x^2 + 3x - 2 = 0$
6. $x^4 - 4x^3 + 6x^2 - 4x - 11 = 0$

7. $x^4 + 4x^3 + 6x^2 + 5x + 3 = 0$
8. $x^4 - 8x^3 + 8x^2 - 8x + 3 = 0$
9. $x^5 - 5x^4 + 4x^3 + 3x^2 - 3x + 2 = 0$
10. $x^5 - 10x^4 + 40x^3 - 50x^2 + 6x - 4 = 0$

6.6 Reciprocal equations. Reciprocal equations are those that remain unaltered when x is changed into its reciprocal. These equations are usually divided into two classes.
(I). In the first class

$$a_0 = a_n, \qquad a_1 = a_{n-1}, \qquad a_2 = a_{n-2}, \cdots, a_{n-1} = a_1,$$
$$a_n = a_0 \qquad (a_0 \neq 0).$$

Here the coefficients of the corresponding terms taken from the beginning and end are equal in magnitude and have the same sign.
(II). In the second class

$$a_0 = -a_n, \quad a_1 = -a_{n-1}, \quad a_2 = -a_{n-2}, \cdots \quad (a_0 \neq 0).$$

Here the coefficients of the corresponding terms taken from the beginning and end are equal in magnitude but different in sign. Notice that in this case when the degree of the equation is even ($n = 2m$), one of the conditions becomes $a_m = -a_m$, or $a_m = 0$. Thus the middle term is absent.

If r is a root of a reciprocal equation, $\frac{1}{r}$ must also be a root, since it is a root of the transformed equation, and the transformed equation is identical with the given equation; hence the roots of a reciprocal equation occur in pairs; r_1, $\frac{1}{r_1}$; r_2, $\frac{1}{r_2}$; etc. When the degree is odd, there must be a root which is its own reciprocal. Hence -1 or $+1$ is then a root, and every root of a reciprocal equation is either its own reciprocal, namely $+1$ or -1, or is paired with its distinct reciprocal.

When the degree is *odd*, we can divide the original equation by the known factor ($x + 1$ or $x - 1$), and the quotient is a reciprocal equation of even degree and of the first class. In equations of the second class of *even* degree, $x^2 - 1$ is a factor, since the equation may be written in the form

$$(x^n - 1) + p_1x(x^{n-2} - 1) + \cdots = 0.$$

By dividing by $x^2 - 1$, this is reducible to a reciprocal equation of the first class of even degree.

6.7 Solution of reciprocal equations. In the previous article we showed that the discussion of every reciprocal equation can be reduced to that of a reciprocal equation of even degree in which the coefficients counting from the beginning and end are equal with the same sign. This is said to be the *standard form.* We now proceed to prove that *a reciprocal equation of the standard form can always be reduced to another equation whose degree is half that of the given reciprocal equation.*

Consider the equation

$$a_0x^{2m} + a_1x^{2m-1} + \cdots + a_mx^m + \cdots + a_1x + a_0 = 0, \quad (a_0 \neq 0).$$

Divide by x^m and unite terms equally distant from the ends. We find

$$a_0\left(x^m + \frac{1}{x^m}\right) + a_1\left(x^{m-1} + \frac{1}{x^{m-1}}\right) + \cdots$$
$$+ a_{m-1}\left(x + \frac{1}{x}\right) + a_m = 0.$$

We have the identity

$$x^{p+1} + \frac{1}{x^{p+1}} \equiv \left(x^p + \frac{1}{x^p}\right)\left(x + \frac{1}{x}\right) - \left(x^{p-1} + \frac{1}{x^{p-1}}\right)$$

for every integral value of $p > 0$. If we put $x + \dfrac{1}{x} = z$ and let $x^p + \dfrac{1}{x^p}$ be represented by V_p, this identity becomes

$$V_{p+1} = V_pz - V_{p-1}.$$

Give p in succession the values 1, 2, 3, etc. We have

$$V_2 = V_1z - V_0 = z^2 - 2,$$
$$V_3 = V_2z - V_1 = z^3 - 3z,$$
$$V_4 = V_3z - V_2 = z^4 - 4z^2 + 2,$$
$$V_5 = V_4z - V_3 = z^5 - 5z^3 + 5z,$$

and so on.

Substituting these values in the above equation, we obtain an

equation of degree m in z. From the values of z those of x can be obtained by solving a quadratic.

Example. Find the roots of the equation

$$x^{10} - 3x^8 + 5x^6 - 5x^4 + 3x^2 - 1 = 0.$$

Dividing by $x^2 - 1$, we have the reciprocal equation

$$x^8 - 2x^6 + 3x^4 - 2x^2 + 1 = 0, \tag{α}$$

or

$$\left(x^4 + \frac{1}{x^4}\right) - 2\left(x^2 + \frac{1}{x^2}\right) + 3 = 0.$$

In terms of z this becomes

$$(z^4 - 4z^2 + 2) - 2(z^2 - 2) + 3 = 0,$$

or $$z^4 - 6z^2 + 9 = 0, \quad \text{or} \quad (z^2 - 3)^2 = 0.$$

Whence

$$z \quad \pm\sqrt{3}, \text{ giving}$$

$$x + \frac{1}{x} = \sqrt{3} \quad \text{and} \quad x + \frac{1}{x} = -\sqrt{3},$$

and

$$x = \frac{\sqrt{3} \pm i}{2} \quad \text{and} \quad x = \frac{-\sqrt{3} \pm i}{2}.$$

These roots are double roots of (α).

Exercises

Solve the following reciprocal equations:

1. $x^4 + x^3 - 4x^2 + x + 1 = 0$
2. $x^5 - 1 = 0; x^6 - 1 = 0$
3. $2x^6 + x^5 - 13x^4 + 13x^2 - x - 2 = 0$
4. $x^6 + x^5 - x^4 - 2x^3 - x^2 + x + 1 = 0$
5. $12x^4 - 56x^3 + 89x^2 - 56x + 12 = 0$
6. $6x^4 - 35x^3 + 62x^2 - 35x + 6 = 0$
7. $16x^8 - 136x^6 + 321x^4 - 136x^2 + 16 = 0$
8. $81x^8 - 1476x^6 + 6886x^4 - 1476x^2 + 81 = 0$
9. $x^{10} + x^8 + x^6 - x^4 - x^2 - 1 = 0$
10. $y^4 + 4y^3 - 3y^2 + 4y + 1 = 0$
11. $y^5 - 4y^4 + y^3 + y^2 - 4y + 1 = 0$

12. $2y^6 - 5y^5 + 4y^4 - 4y^2 + 5y - 2 = 0$
13. $6x^4 + 5x^3 - 38x^2 + 5x + 6 = 0$
14. $x^{10} - 29x^8 + 226x^6 - 226x^4 + 29x^2 - 1 = 0$
15. $x^{10} - 13x^8 + 50x^6 - 50x^4 + 13x^2 - 1 = 0$
16. Solve $(1 + x)^5 = a(1 + x^5)$ for $a = 1, 5, 16, 45, 1/31$
17. Find the quadratic factors of $x^6 + 1 = 0$
18. Solve $(1 + x)^4 = a(1 + x^4)$.
19. $3x^4 + 10x^3 + 6x^2 + 10x + 3 = 0$
20. $2x^4 + 9x^3 + 14x^2 + 9x + 2 = 0$
21. $3x^4 + 13x^3 + 16x^2 + 13x + 3 = 0$
22. $x^8 - 2x^6 + 3x^4 - 2x^2 + 1 = 0$
23. $2x^4 + 5x^3 + 4x^2 + 5x + 2 = 0$
24. $2x^6 + 9x^5 + 16x^4 + 18x^3 + 16x^2 + 9x + 2 = 0$
25. $x^6 + x^4 + x^2 + 1 = 0$
26. $x^8 - 12x^6 + 6x^4 - 12x^2 + 1 = 0$
27. $36x^6 + 132x^5 - 47x^4 - 386x^3 - 47x^2 + 132x + 36 = 0$

Chapter VII

CUBIC AND QUARTIC EQUATIONS

7.1 The cubic equation. Consider the general cubic with real coefficients.

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0 \qquad (a_0 > 0) \tag{1}$$

To suppose that a_0 is positive is no real restriction, for if a_0 were zero, the equation would be at most quadratic and for $a_0 < 0$ one could change sign throughout. Dividing by a_0, (1) can be written in the form

$$x^3 + bx^2 + cx + d = 0 \tag{2}$$

where

$$b = a_1/a_0, \qquad c = a_2/a_0, \qquad d = a_3/a_0.$$

Making the substitution $x = y - b/3$, we obtain the *reduced cubic equation* (with real coefficients) in which the second-degree term is absent:

$$y^3 + py + q = 0, \tag{3}$$

where

$$p = c - \frac{b^2}{3} \quad \text{and} \quad q = d - \frac{bc}{3} + \frac{2b^3}{27}.$$

If b had been zero at the start, this transformation would merely replace x by y, and the equation (2) would have been in reduced form.

If in the reduced cubic (3), $p = q = 0$, then the three roots of (3) coincide, and (2) is a perfect cube. In this case no further discussion is necessary.

If $p = 0$, but $q \neq 0$, then (3) is in the binomial form, and may be solved if we find the three cube roots of $-q$. This case also is thus readily disposed of.

If $p \neq 0$, but $q = 0$, equation (3) has for roots, $0, \pm\sqrt{-p}$. The three roots of (3) and, therefore, also of (2) will then be in arithmetic progression.

There is one other special case that may be handled before developing any general formula for the solution of (3). This is the case in which (2) and hence also (3) has a repeated root, say r. Since in (3) the sum of the roots is zero, the roots of (3) will be r, r, $-2r$, and (3) may be written in the form

$$y^3 - 3r^2y + 2r^3 = 0.$$

In this case $p = -3r^2$, $q = 2r^3$, so that

$$\Delta = -4p^3 - 27q^2 = 0.$$

We shall encounter this expression Δ again later. In this special case, $r = -3q/2p$, and the roots are of the form, $-3q/2p$, $-3q/2p$, $3q/p$.

Save for these special cases, it is no restriction upon the problem to suppose that

$$p \neq 0, \qquad q \neq 0, \qquad \Delta = -4p^3 - 27q^2 \neq 0.$$

Let us now return to (3) and introduce two new variables A and B, and their respective cube roots, u, v, where

$$u^3 = A \text{ and } v^3 = B.$$

Since a number has three cube roots, u and v each has three values. Put

$$= u + v; \tag{4}$$

then y has nine values. Now a cubic has only three roots. We must place some restrictions on u and v so that y will have three and only three values.

Substitute (4) in (3). We have

$$(u + v)^3 + p(u + v) + q = 0,$$

or (5)

$$u^3 + v^3 + (3uv + p)(u + v) + q = 0.$$

Now let us notably simplify equation (5) by imposing the condition on u and v that

$$3uv + p = 0. \tag{6}$$

Then

$$u^3v^3 = -\frac{p^3}{27}. \tag{7}$$

Because of (6), equation (5) reduces to

$$u^3 + v^3 = -q. \tag{8}$$

From the relations (7) and (8), u^3 and v^3 may be considered as the two roots, A and B, of the quadratic equation

$$z^2 + qz - \frac{p^3}{27} = 0. \tag{9}$$

The relatively trivial cases, $p = 0$, or $q = 0$, or $4p^3 + 27q^2 = 0$, have been disposed of already. They would otherwise call for consideration here. Solving this quadratic equation, we have

$$z = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}.$$

Hence we may set

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = A,$$

and (10)

$$v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = B.$$

Then

$$u = \sqrt[3]{A};\, \omega\sqrt[3]{A};\, \omega^2\sqrt[3]{A},$$

and (11)

$$v = \sqrt[3]{B};\, \omega\sqrt[3]{B};\, \omega^2\sqrt[3]{B},$$

where

$$\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \quad \text{and} \quad \omega^2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

From (6) the values of u and v in (4) must be paired in such a way that $uv = -p/3$. Exactly three pairs of values in (11) satisfy this condition: these are

$$\sqrt[3]{A},\, \sqrt[3]{B}; \qquad \omega\sqrt[3]{A},\, \omega^2\sqrt[3]{B}; \qquad \omega^2\sqrt[3]{A},\, \omega\sqrt[3]{B}.$$

Substituting these values of u and v in (4), we obtain the three roots of (3), namely

$$y_1 = \sqrt[3]{A} + \sqrt[3]{B};\; y_2 = \omega\sqrt[3]{A} + \omega^2\sqrt[3]{B};\; y_3 = \omega^2\sqrt[3]{A} + \omega\sqrt[3]{B}. \tag{12}$$

This solution offers no difficulties if A and B are real. The case where A and B are imaginary is taken up later in this chapter.

Example 1. Solve $y^3 + 6y + 20 = 0$.
We have

$$p = 6, \quad q = 20$$

whence

$$z^2 + qz - \frac{p^3}{27} = 0 \quad \text{becomes} \quad z^2 + 20z - 8 = 0.$$

Then

$$z_1 = u^3 = -10 + 6\sqrt{3} = 0.39230 \text{ and } u = 0.732,$$

and

$$z_2 = v^3 = -10 - 6\sqrt{3} = -20.39230 \text{ and } v = -2.732.$$

One can verify that $u = -1 + \sqrt{3}$ and $v = -1 - \sqrt{3}$ serve as real values. The three values of u and v, respectively, are as follows:

$$u = \quad 0.732, \quad 0.732\omega, \quad 0.732\omega^2;$$
$$v = -2.732, \ -2.732\omega, \ -2.732\omega^2.$$

Whence

$$y_1 = -2; \qquad y_2 = 1 + 3i; \qquad y_3 = 1 - 3i.$$

Example 2. Solve $y^3 - 6y - 6 = 0$.
We have

$$p = -6, \qquad q = -6;$$

so

$$z^2 + qz - \frac{p^3}{27} = 0 \quad \text{becomes} \quad z^2 - 6z + 8 = 0.$$

Then

$$z_1 = u^3 = 4 \qquad \text{and} \qquad u = \sqrt[3]{4} = 1.5874;$$
$$z_2 = v^3 = 2 \qquad \text{and} \qquad v = \sqrt[3]{2} = 1.2599.$$

Therefore we find

$$y = \sqrt[3]{4} + \sqrt[3]{2}; \quad \omega\sqrt[3]{4} + \omega^2\sqrt[3]{2}; \quad \omega^2\sqrt[3]{4} + \omega\sqrt[3]{2}$$
$$= 2.8473; \quad -1.4237 + 0.2836i; \quad -1.4237 - 0.2836i.$$

Exercises

Solve the following cubic equations:

1. $y^3 - 3y + 2 = 0$

2. $y^3 - 9y - 28 = 0$
3. $y^3 - 18y - 35 = 0$
4. $y^3 - 12y - 16 = 0$
5. $y^3 - 27y + 54 = 0$
6. $x^3 - 12x^2 + 30x - 27 = 0$
7. $x^3 - 9x^2 + 24x - 20 = 0$
8. $y^3 - 3y - 18 = 0$
9. $y^3 + 3y - 14 = 0$
10. $y^3 + 3y - 36 = 0$
11. $y^3 + 18y - 215 = 0$
12. $y^3 + 6y - 20 = 0$
13. $x^3 + 3x^2 - 3x - 45 = 0$
14. $x^3 - 6x^2 + 10x - 8 = 0$
15. $x^3 - 9x^2 + 28x - 30 = 0$
16. $x^3 - 6x^2 + 13x - 10 = 0$
17. $x^3 - 6x^2 + 3x - 18 = 0$
18. $y^3 + 15y - 124 = 0$
19. $x^3 - 6x^2 + 33x - 392 = 0$
20. $x^3 - 15x^2 + 93x - 196 = 0$

7.2 Discriminant. The discriminant of any equation with real coefficients in which the coefficient of the highest power of the unknown is plus one, is defined to be the product of the squares of the differences of the roots. For the reduced cubic the discriminant Δ is

$$\Delta = (y_1 - y_2)^2(y_2 - y_3)^2(y_3 - y_1)^2 = -4p^3 - 27q^2.$$

This may be proved as follows. For convenience we repeat equations (12).

$$y_1 = \sqrt[3]{A} + \sqrt[3]{B};\ y_2 = \omega\sqrt[3]{A} + \omega^2\sqrt[3]{B};\ y_3 = \omega^2\sqrt[3]{A} + \omega\sqrt[3]{B}. \quad (12)$$

Multiply these expressions for y_1, y_2, y_3 by 1, ω^2, ω and add.
Multiply these expressions for y_1, y_2, y_3 by 1, ω, ω^2 and add.
We obtain

$$\frac{1}{3}(y_1 + \omega^2 y_2 + \omega y_3) = \sqrt[3]{A} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad \text{(a)}$$

$$\frac{1}{3}(y_1 + \omega y_2 + \omega^2 y_3) = \sqrt[3]{B} = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \quad \text{(b)}$$

Take the difference of the cubes of these expressions. We have

$$\frac{1}{54}[(y_1 + \omega^2 y_2 + \omega y_3)^3 - (y_1 + \omega y_2 + \omega^2 y_3)^3] = \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}; \qquad \text{(c)}$$

also

$$\frac{1}{54}[(y_1 + \omega^2 y_2 + \omega y_3)^3 - (y_1 + \omega y_2 + \omega^2 y_3)^3]$$
$$= \frac{\sqrt{-3}}{18}(y_1 - y_2)(y_2 - y_3)(y_3 - y_1).$$

Squaring and clearing of fractions, we find

$$\Delta = (y_1 - y_2)^2(y_2 - y_3)^2(y_3 - y_1)^2 = -4p^3 - 27q^2. \qquad (13)$$

From (a), (b) and (c) we see that the irrationalities appearing in the solution of the cubic can be expressed rationally in terms of the roots. One may readily verify the following:

$$\Delta = \begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix}^2 = [x_2x_3(x_2 - x_3) + x_3x_1(x_3 - x_1) + x_1x_2(x_1 - x_2)]^2$$
$$= [x_1^2(x_2 - x_3) + x_2^2(x_3 - x_1) + x_3^2(x_1 - x_2)]^2$$
$$= \tfrac{1}{9}[(x_2 - x_3)^3 + (x_3 - x_1)^3 + (x_1 - x_2)^3]^2.$$

We observe also that the discriminant is expressed rationally in terms of the coefficients of the reduced cubic.

The discriminant of the general cubic $x^3 + bx^2 + cx + d = 0$, is equal to the discriminant of the reduced cubic $y^3 + py + q = 0$. This may be proved as follows.

Since the reduced cubic is derived from the general cubic by the substitution $x = y - \frac{b}{3}$, the roots of the general cubic are

$$x_1 = y_1 - \frac{b}{3}; \qquad x_2 = y_2 - \frac{b}{3}; \qquad x_3 = y_3 - \frac{b}{3}.$$

Whence

$$x_1 - x_2 = y_1 - y_2; \quad x_1 - x_3 = y_1 - y_3; \quad x_2 - x_3 = y_2 - y_3.$$

Therefore

$$(x_1 - x_2)^2(x_2 - x_3)^2(x_3 - x_1)^2 = (y_1 - y_2)^2(y_2 - y_3)^2(y_3 - y_1)^2.$$

Replacing p and q by their values, namely

$$p = c - \frac{b^2}{3}, \qquad q = d - \frac{bc}{3} + \frac{2b^3}{27},$$

we have

$$\Delta = -4p^3 - 27q^2 = 18bcd - 4b^3d + b^2c^2 - 4c^3 - 27d^2.$$

It may be noted also that the discriminant Δ of $x^3 + bx^2 + cx + d$ is -27 times the discriminant of the quadratic

$$z^2 + qz - \frac{p^3}{27} = 0 \tag{9}$$

used in solving the cubic.

7.3. Nature of the roots of a cubic.

I. If $\Delta = 0$ *all roots are real. Two roots are equal.*

If $p = 0$ or $q = 0$, we see from (13) that the other is also, in which case from (10) and (12) we see that $y_1 = y_2 = y_3 = 0$.

If $q \neq 0$, then from (10) $A = B = -\frac{q}{2}$, whence from (12)

$$y_2 = y_3 = (\omega + \omega^2)\sqrt[3]{-\frac{q}{2}} = -\sqrt[3]{-\frac{q}{2}}$$

and

$$y_1 = 2\sqrt[3]{-\frac{q}{2}},$$

or in rational form, as discussed earlier,

$$y_1 = \frac{3q}{p}, \qquad y_2 = y_3 = -\frac{3q}{2p}.$$

II. If $\Delta < 0$, *one root is real and two are imaginary.*

From (10), A and B are real and distinct. Hence from (12), y_1 is real and y_2 and y_3 are conjugate imaginary numbers.

III. If $\Delta > 0$, *all three roots are real and distinct.*

From (10), A and B are conjugate imaginary numbers.
Let $A = a + bi$ and $B = a - bi$.
Then $\sqrt[3]{A}$ and $\sqrt[3]{B}$ will also be conjugate imaginary numbers, for suitable choices of the cube roots.
Let

$$\sqrt[3]{A} = c + di \qquad \text{and} \qquad \sqrt[3]{B} = c - di \qquad (d \neq 0).$$

The three roots will be

$$y_1 = \quad (c + di) + \quad (c - di) = 2c$$

$$y_2 = \omega\,(c + di) + \omega^2(c - di) = -c - d\sqrt{3} \qquad (d \neq 0)$$
$$y_3 = \omega^2(c + di) + \omega\,(c - di) = -c + d\sqrt{3}.$$

If $\Delta > 0$, the solution is usually obtained by trigonometric means. The formulas (12) present the roots in a form involving cube roots of imaginaries. This is called the *irreducible case*, since it can be shown that the cube root of a general complex number cannot be expressed in the form $c + di$ where c and d involve only real radicals.*

Exercises

For the following equations, compute the discriminant and determine the nature of the roots:

1. $y^3 + 3y + 3 = 0$
2. $y^3 - 6y + 3 = 0$
3. $y^3 - 3y + 1 = 0$
4. $y^3 - 3y + 3 = 0$
5. $y^3 - 3y + 2 = 0$
6. $y^3 - 12y + 16 = 0$
7. $y^3 - 27y + 54 = 0$
8. $x^3 + x^2 - 5x + 3 = 0$
9. $x^3 + x^2 + 2x + 3 = 0$
10. $x^3 + 3x^2 + 2x + 1 = 0$
11. Compute the discriminant for each cubic in the previous set of exercises.
12. Show that for $x^3 + bx^2 + cx + d$, $\Delta = 0$, if and only if the two quadratic expressions $3x^2 + 2bx + c$ and $bx^2 + 2cx + 3d$ have a common factor containing x.

7.4 Trigonometric solution of a cubic equation with $\Delta > 0$. This is the so-called irreducible case. It is of interest that this is the case in which all of the roots are real. They can be computed with the help of a table of cosines.

From trigonometry we have the identity

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta$$

so that

$$\cos^3\theta - \tfrac{3}{4}\cos\theta - \tfrac{1}{4}\cos 3\theta = 0. \tag{14}$$

In

$$y^3 + py + q = 0, \quad \text{put} \quad y = z/n; \quad \text{then}$$
$$z^3 + pn^2z + n^3q = 0. \tag{15}$$

* See L. E. Dickson's *Elementary Theory of Equations*, pp. 35-36, John Wiley & Sons, 1914.

Equations (14) and (15) will be identical if

$$z = \cos\theta; \qquad pn^2 = -\tfrac{3}{4}; \qquad n^3q = -\tfrac{1}{4}\cos 3\theta.$$

Whence

$$n = \sqrt{-3/(4p)}$$

and

$$\cos 3\theta = -4q\left(-\frac{3}{4p}\right)^{\frac{3}{2}} = -\frac{q}{2}\left(-\frac{27}{p^3}\right)^{\frac{1}{2}}. \tag{16}$$

These equations can always be solved if p is negative, and

$$\left|\frac{q}{2}\left(-\frac{27}{p^3}\right)^{\frac{1}{2}}\right| < 1.$$

This last condition reduces to $-4p^3 - 27q^2 = \Delta > 0$, and so is satisfied in the cases under consideration.

If θ is the smallest angle satisfying (16), then the values

$$\theta + 120^\circ \quad \text{and} \quad \theta + 240^\circ$$

also satisfy it, so that the roots of the equation

$$y^3 + py + q = 0$$

are

$$\frac{1}{n}\cos\theta, \qquad \frac{1}{n}\cos(\theta + 120^\circ), \qquad \frac{1}{n}\cos(\theta + 240^\circ)$$

correct to a number of decimal places depending on the tables used.

Example. Solve

$$y^3 - 3y + 1 = 0.$$

We have

$$p = -3, \quad q = 1.$$

Then

$$n = \sqrt{-\frac{3}{-12}} = \frac{1}{2}$$

and

$$\cos 3\theta = -4n^3q = -\tfrac{1}{2} = \cos 120^\circ.$$

Therefore $\theta = 40^\circ$. Hence the solutions are

$$2\cos 40^\circ, \qquad 2\cos 160^\circ, \qquad 2\cos 280^\circ$$

or

$$1.53208, \qquad -1.87938, \qquad 0.34730.$$

Exercises

Solve the following cubic equations:

1. $y^3 - y + \frac{1}{3} = 0$	5. $y^3 - 12y - 10 = 0$
2. $y^3 - 6y - 4 = 0$	6. $x^3 + 9x^2 + 24x + 19 = 0$
3. $y^3 - 3y - 1 = 0$	7. $3x^3 + 3x^2 - 3x - 2 = 0$
4. $x^3 - 6x^2 - x + 30 = 0$	8. $x^3 + 9x^2 + 18x - 10 = 0$

7.5 Graphical solution of the cubic. Every cubic, with real coefficients, has one real root. Let the cubic be

$$f(x) \equiv x^3 + ax^2 + bx + c = 0. \qquad (\Delta < 0)$$

Let the real root be represented in Fig. 20 by OA. Draw any secant line through A cutting the cubic at B and C. Let D be

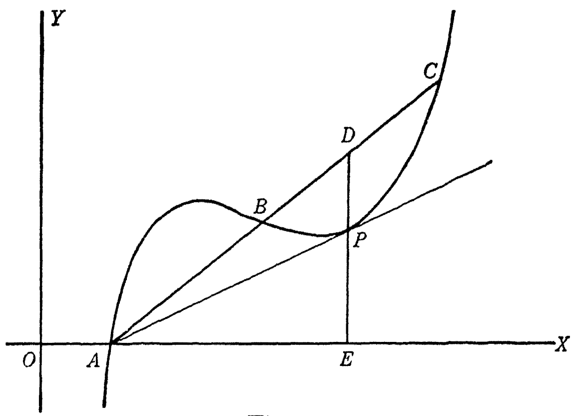


Fig. 20

the midpoint of BC. Let the ordinate DE cut the cubic in P. Then AP is tangent to the cubic and the imaginary roots $u \pm vi$ are such that $u = OE$ and $v = \sqrt{\tan EAP}$.

Proof. Let the roots of the cubic be r, $u \pm vi$ where r, u, v are real. Then

$$f(x) = (x - r)(x^2 - 2ux + u^2 + v^2) = 0.$$

Let $y = m(x - r)$ represent any secant through A.
The x coordinates of the points of intersection of the secant line with the cubic are given by

$x^2 - 2ux + u^2 + v^2 - m = 0$; that is, by $x = u \pm \sqrt{m - v^2}$.
The secant will be a tangent if $m = v^2$.
The perpendicular from C upon ED is $\sqrt{m - v^2}$, which reduces to v if $m = 2v^2 = 2 \tan EAP$.

7.6 Solution of the quartic. It is convenient for the present purpose to write the quartic in the form

$$x^4 + 2px^3 + qx^2 + rx + s = 0. \tag{17}$$

This may be written

$$x^4 + 2px^3 = -qx^2 - rx - s. \tag{18}$$

The plan of solution presented here depends upon the possibility of adding to both sides of this equation an expression of the second degree in x such that the left member will be the square of a second degree expression in x and the right member will be the square of a first degree expression in x. Since for every value of u,

$$\left(x^2 + px + \frac{u}{2}\right)^2 = x^4 + 2px^3 + (p^2 + u)x^2 + pux + \frac{u^2}{4},$$

we see that the expression which we may add to both sides is

$$(p^2 + u)x^2 + pux + u^2$$

Adding this expression to both sides of (18), we have

$$\left(x^2 + px + \frac{u}{2}\right)^2 = (p^2 + u - q)x^2 + (pu - r)x + \frac{u^2}{4} - s. \tag{19}$$

The left member of (19) is a perfect square for every value of u. It remains to determine u so that the right member is a perfect square. The right member will be a perfect square if u satisfies the following condition derived from the condition for equal roots of a quadratic:

$$(pu - r)^2 - 4(p^2 + u - q)\left(\frac{u^2}{4} - s\right) = 0,$$

or

$$u^3 - qu^2 + (2pr - 4s)u + (4qs - 4p^2s - r^2) = 0. \tag{20}$$

This is called a reducing cubic equation for the quartic (17), for if we can solve (20) for u, then for each such solution, u, (19) and hence (17) can be solved.

Let u_1, u_2, u_3 be the roots of (20). Let one of these roots, say u_1, be substituted for u in (19). Then (19) can be written as follows:

$$\left(x^2 + px + \frac{u_1}{2}\right)^2 = (a_1x + b_1)^2.$$

Hence

$$x^2 + px + \frac{u_1}{2} = a_1x + b_1 \tag{21}$$

and

$$x^2 + px + \frac{u_1}{2} = -a_1x - b_1. \tag{22}$$

The four roots x_1, x_2, x_3, x_4 of (21) and (22) are the roots of (17). The quartic cannot have more than four roots. It follows that the four roots obtained from

$$\left.\begin{aligned} x^2 + px + \frac{u_2}{2} &= a_2x + b_2 \\ x^2 + px + \frac{u_2}{2} &= -a_2x - b_2 \end{aligned}\right\} \tag{23}$$

or the four roots obtained from

$$\left.\begin{aligned} x^2 + px + \frac{u_3}{2} &= a_3x + b_3 \\ x^2 + px + \frac{u_3}{2} &= -a_3x - b_3 \end{aligned}\right\} \tag{24}$$

will be the same as those obtained from (21) and (22). The roots, however, will be paired differently.

Example. Solve $x^4 + 6x^3 + 15x^2 + 26x + 24 = 0$.

Equation (20) becomes $u^3 - 15u^2 + 60u - 100 = 0$, whence

$$u = 10, \quad \tfrac{1}{2}(5 \pm i\sqrt{15}).$$

Let us use $u = 10$. Equations (21) and (22) become

$$x^2 + 3x + 5 = 2x + 1 \quad \text{and} \quad x^2 + 3x + 5 = -2x - 1,$$

whence

$$x^2 + x + 4 = 0 \qquad \text{and} \quad x^2 + 5x + 6 = 0.$$

Therefore

$$x = -\frac{1}{2} \pm i\frac{\sqrt{15}}{2} \quad \text{and} \quad x = -2, -3.$$

Let the given quartic be $z^4 + az^3 + bz^2 + cz + d = 0$. The term in z^3 can be removed by the substitution $z = x - \frac{a}{4}$. Verify that the transformed quartic is $x^4 + qx^2 + rx + s = 0$ where

$$q = b - \frac{3}{8}a^2; \quad r = c - \frac{ab}{2} + \frac{a^3}{8}; \quad s = d - \frac{1}{4}ac + \frac{1}{16}a^2b - \frac{3}{256}a^4.$$

The equations (19) (20) (21) (22) (23) (24) now serve for the solution of the transformed equation by putting $p = 0$.

7.7 Roots of the resolvent cubic in terms of roots of the quartic. Put

$$V_1 = x_1x_2 + x_3x_4; \quad V_2 = x_1x_3 + x_2x_4; \quad V_3 = x_1x_4 + x_2x_3. \tag{25}$$

From (17) we have

$$x_1 + x_2 + x_3 + x_4 = -2p; \; x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 = -r;$$
$$x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = q; \; x_1x_2x_3x_4 = s.$$

Whence we are able to derive the following relations:

$$\begin{aligned}
V_1 + V_2 + V_3 &= x_1x_2 + x_3x_4 + x_1x_3 + x_2x_4 + x_1x_4 + x_2x_3 = q \\
V_1V_2 + V_1V_3 + V_2V_3 &= (x_1x_2 + x_3x_4)(x_1x_3 + x_2x_4) + \cdots \\
&= (x_1 + x_2 + x_3 + x_4)(x_1x_2x_3 + x_1x_2x_4 + \\
&\qquad x_1x_3x_4 + x_2x_3x_4) - 4x_1x_2x_3x_4 \\
&= 2pr - 4s \\
V_1V_2V_3 &= (x_1x_2 + x_3x_4)(x_1x_3 + x_2x_4)(x_1x_4 + x_2x_3) \\
&= (x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4)^2 \\
&\quad + x_1x_2x_3x_4[(x_1 + x_2 + x_3 + x_4)^2 - 4(x_1x_2 + \\
&\qquad x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4)] \\
&= r^2 + s(4p^2 - 4q) = r^2 + 4p^2s - 4qs.
\end{aligned}$$

Hence the values of V_1, V_2, V_3 given in (25) are the roots of the cubic (20).

7.8 Graphical construction of the roots of a quartic. Let us consider the quartic with two real and two imaginary roots:

$$a \pm b, \quad \alpha \pm \beta i$$

$$y = f(x) \equiv x^4 - 2(a + \alpha)x^3 + (a^2 - b^2 + \alpha^2 + \beta^2 + 4a\alpha)x^2$$
$$-2(a\alpha^2 + a\beta^2 + a^2\alpha - b^2\alpha)x + (a^2 - b^2)(\alpha^2 + \beta^2) = 0.$$

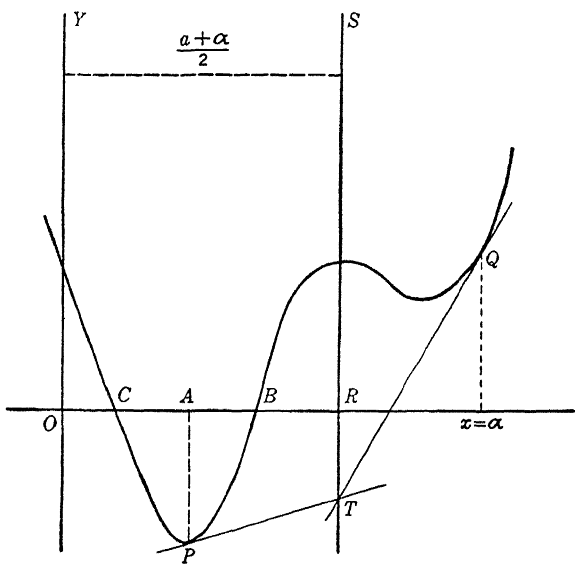


Fig. 21

Let

$$OA = a, \quad CA = AB = b.$$

Then the real roots are represented by the points C and B in Fig. 21. If

$$x = a$$
$$y = -b^2(a - \alpha)^2 - b^2\beta^2,$$
$$\frac{dy}{dx} = -2b^2(a - \alpha).$$

The equation of the tangent PT at the point P whose abscissa is a is

$$y + b^2(a - \alpha)^2 + b^2\beta^2 = -2b^2(a - \alpha)(x - a).$$

This tangent intersects the line $x = \frac{1}{2}(a + \alpha)$ where $y = -b^2\beta^2$.
If

$$x = \alpha, \qquad y = \beta^2(a - \alpha)^2 - b^2\beta^2, \qquad \frac{dy}{dx} = -2\beta^2(a - \alpha).$$

The equation of the tangent QT at the point Q whose abscissa is α is

$$y - \beta^2(a - \alpha)^2 + b^2\beta^2 = -2\beta^2(a - \alpha)(x - \alpha).$$

This line intersects the line $x = \frac{1}{2}(a + \alpha)$ where $y = -b^2\beta^2$.

To obtain the real roots carefully draw the graph. Measure OC and OB. This gives approximate values of the real roots. Find A the mid point of CB. The ordinate at A cuts the curve in P. Draw the tangent at P. This cuts the line $x = \frac{1}{2}(a + \alpha)$ at T whose ordinate is $-b^2\beta^2$. Through T draw a tangent TQ. This touches the curve at some point Q. The abscissa of Q will be the desired value α. $RT = -b^2\beta^2$. b is known. Then β can be found by graphical means. While of theoretical interest, this construction is of little value in finding numerical values of the imaginary roots.

Exercises

Solve the following quartics:

1. $x^4 + 6x^3 + 12x^2 + 10x + 3 = 0$
2. $x^4 - 10x^2 - 20x - 16 = 0$
3. $x^4 + x^2 + 4x - 3 = 0$
4. $x^4 - 6x^3 + 12x^2 - 20x - 12 = 0$
5. $x^4 + x^3 - x^2 - 7x - 6 = 0$
6. $x^4 + 2x^3 - 12x^2 - 10x + 3 = 0$
7. $x^4 + 4x^3 - 6x^2 + 4x + 8 = 0$
8. $x^4 + 6x^3 + 7x^2 - 6x - 8 = 0$
9. $x^4 + 6x^3 + 3x^2 + 6x - 8 = 0$
10. $x^4 - 17x^2 + 20x + 12 = 0$
11. $x^4 - 38x^2 + 108x - 80 = 0$
12. $x^4 - 8x^2 - 8x + 15 = 0$
13. $x^4 - 32x^2 + 72x - 32 = 0$
14. $x^4 - 20x^2 + 48x - 32 = 0$

15. $x^4 - 22x^2 - 72x - 72 = 0$
16. $x^4 - 10x^2 - 72x - 72 = 0$
17. $x^4 - 30x^2 + 60x - 16 = 0$
18. $x^4 + 32x + 48 = 0$
19. $x^4 + 6x^3 + 4x^2 - 54x - 117 = 0$
20. $x^4 + 6x^3 - 2x^2 - 96x - 224 = 0$
21. $x^4 + 2x^3 - 2x - 1 = 0$
22. $x^4 + 2x^3 - 7x^2 - 8x + 12 = 0$

Chapter VIII

BOUNDS FOR THE ROOTS OF AN EQUATION

8.1 Introduction. The problem before us is to find the roots of equations. It is advantageous to be able to narrow the region within which these roots must be sought. In Chapter IV we already have proved two theorems which help us in this matter. For reference purposes it will be helpful to repeat these theorems here, that we may have in one chapter all of our theorems with respect to the bounds of roots of equations. The theorems bear here the same numbers that they have in Chapter IV.

Theorem 1. *If in the polynomial, with real coefficients,*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n \quad (a_0 > 0)$$

the value $\left|\dfrac{a_k}{a_0}\right| + 1$, *or any greater value, be substituted for* x, *where* a_k *is that one of the coefficients* $a_1, a_2, \cdots, a_n$ *whose numerical value is greatest, the term containing the highest power of* x *will exceed, numerically, the sum of all the terms which follow.*

In terms of the roots of the equation $f(x) = 0$, this theorem means that no root of the equation can exceed $\left|\dfrac{a_k}{a_0}\right| + 1$.

Theorem 2. *If in the polynomial, with real coefficients,*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n \quad (a_n \neq 0)$$

the value $\dfrac{|a_n|}{|a_n| + |a_k|}$, *or any smaller positive value, be substituted for* x, *where* a_k *is that one of the coefficients* $a_0, a_1, a_2, \cdots, a_{n-1}$ *whose numerical value is greatest, the term* a_n *will be numerically greater than the sum of all the others.*

In terms of the roots of the equation $f(x) = 0$, this theorem means that no positive root of the equation can be less than $\dfrac{|a_n|}{|a_n| + |a_k|}$. We will proceed in this chapter to find other bounds of the roots of equations.

Theorem 3. *In any algebraic equation, with real coefficients,*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 > 0)$$

if the first negative coefficient is preceded by r coefficients which are positive or zero, and if the numerically greatest negative coefficient is a_k, then each real root is less than $1 + \sqrt[r]{|a_k|/a_0}$.

We desire to find a value of x for which, and for any larger value of x, $f(x) > 0$. Omit the terms $a_1x^{n-1} + \cdots + a_{r-1}x^{n-r+1}$, which are positive or zero for positive values of x, and replace each coefficient after a_{r-1} by $-|a_k|$; then certainly

$$f(x) \geqq a_0x^n - |a_k|(x^{n-r} + x^{n-r-1} + \cdots + x + 1).$$

Hence $f(x)$ will be positive if

$$a_0x^n - |a_k|(x^{n-r} + x^{n-r-1} + \cdots + x + 1) > 0. \qquad (1)$$

But

$$x^{n-r} + x^{n-r-1} + \cdots + x + 1 = \frac{x^{n-r+1} - 1}{x - 1}, \qquad (x \neq 1).$$

Hence for $x > 1$, the inequality (1) is satisfied if

$$a_0x^n > |a_k| \frac{x^{n-r+1}}{x-1},$$

or

$$a_0x^{r-1}(x - 1) > |a_k|$$

which inequality again is satisfied if

$$a_0(x - 1)^{r-1}(x - 1) \geq |a_k|$$

since obviously $x^{r-1} > (x - 1)^{r-1}$. We consider, therefore, finally

$$a_0(x - 1)^r \geq |a_k|$$

or

$$x \geq 1 + \sqrt[r]{\frac{|a_k|}{a_0}}.$$

Exercises

Find an upper bound of the roots of the following equations:

1. $x^4 + 2x^3 - 21x^2 - 22x + 40 = 0$
2. $x^4 + 4x^3 - 34x^2 - 76x + 105 = 0$

3. $x^4 + x^3 - 63x^2 - 64x - 64 = 0$
4. $x^4 - 143x^2 - 144 = 0$
5. $x^4 - 106x^2 + 600 = 0$
6. $x^4 - 73x^2 + 576 = 0$
7. $x^4 - 27x^2 + 14x + 120 = 0$
8. $x^5 - 29x^4 - 31x^3 + 31x^2 - 32x + 60 = 0$
9. $x^4 + 2x^3 - 27x^2 - 28x - 60 = 0$
10. $x^6 + x^5 - x^4 - 2x^3 - x^2 + x + 1 = 0$
11. $x^6 - 103x^4 + 296x^2 + 400 = 0$
12. $x^7 + x^6 - 103x^5 - 103x^4 + 296x^3 + 296x^2 + 400x + 400 = 0$
13. $x^4 + 6x^3 + 4x^2 - 54x - 117 = 0$
14. $x^4 - 10x^2 - 20x - 16 = 0$
15. $x^4 + x^3 - x^2 - 7x - 6 = 0$
16. $x^4 + 2x^3 - 12x^2 - 10x + 3 = 0$
17. $x^4 + 6x^3 + 7x^2 - 6x - 8 = 0$
18. $x^4 - 17x^2 + 20x + 12 = 0$
19. $x^4 - 18x^2 + 32x - 15 = 0$
20. $2x^4 + 3x^3 + 4x^2 - 5x - 4 = 0$
21. $2x^6 + x^5 - 13x^4 + 13x^2 - x - 2 = 0$
22. $6x^4 + 5x^3 - 38x^2 + 5x + 6 = 0$
23. $x^7 + 3x^4 - 16x^3 - 16x^2 - 16x + 48 = 0$
24. $x^6 + x^4 - 27x^3 - 26x^2 + 3x + 144 = 0$
25. $2x^6 + x^4 - 54x^3 + x^2 + 20x - 150 = 0$
26. $2x^5 + x^3 - 128x^2 - 40x + 96 = 0$
27. $x^6 - 256x^2 + 18x - 72 = 0$
28. $6x^7 - 486x^3 + 30x^2 - 78x - 36 = 0$
29. $3x^7 + 2x^5 + x^4 - 243x^3 + 300x^2 + 50x + 180 = 0$
30. $8x^6 - 216x^3 + 21x^2 - 43x - 60 = 0$

8.2 Theorem 4: *If in any rational integral algebraic equation, with real coefficients, the numerical value of each negative coefficient be divided by the sum of all the positive coefficients which precede it, the greatest of the fractions so formed, increased by one, is an upper bound to the roots of the equation.*

To fix the ideas, let the equation be

$$f(x) \equiv a_0x^n + a_1x^{n-1} - a_2x^{n-2} + \cdots - a_rx^{n-r} + \cdots + a_n = 0 \qquad (a_0 > 0)$$

in which, for illustration, we regard the third coefficient as negative, and we have a general negative coefficient, namely $-a_r$.

Since

$$\frac{x^m - 1}{x - 1} = x^{m-1} + x^{m-2} + \cdots + x + 1$$

we have

$$a_{n-m}x^m = a_{n-m}(x - 1)(x^{m-1} + x^{m-2} + \cdots + x + 1) + a_{n-m}.$$

Transforming the *positive* terms in $f(x)$ by means of this formula, the polynomial $f(x)$ becomes

$$\begin{aligned}
&a_0(x - 1)x^{n-1} + a_0(x - 1)x^{n-2} + a_0(x - 1)x^{n-3} \\
&\qquad\qquad\qquad + \cdots + a_0(x - 1)x^{n-r} + \cdots + a_0 \\
&\qquad + a_1(x - 1)x^{n-2} + a_1(x - 1)^{n-3} \\
&\qquad\qquad\qquad + \cdots + a_1(x - 1)x^{n-r} + \cdots + a_1 \\
&\qquad\qquad - a_2x^{n-2} \\
&\qquad\qquad\qquad\qquad + \cdots\cdots\cdots \\
&\qquad\qquad\qquad\qquad - a_rx^{n-r}, \\
&\qquad\qquad\qquad\qquad\qquad + \cdots\cdots \\
&\qquad\qquad\qquad\qquad\qquad + a_n,
\end{aligned}$$

where the horizontal lines correspond to the successive terms of $f(x)$.

In those columns in which a coefficient of the form $-a_i$ occurs, only one such negative value can occur.

We now regard the sum of elements in a vertical column of this expression as a term in the polynomial; the successive coefficients of $x^{n-1}, x^{n-2}, \cdots, x^{n-r}, \cdots$ being

$$a_0(x - 1), (a_0 + a_1)(x - 1) - a_2, \cdots,$$
$$(a_0 + a_1 + \cdots + a_{r-1})(x - 1) - a_r, \cdots.$$

If x is given a value large enough so that the sum of the terms in each column is positive, then $f(x)$ is positive. Those columns in which every a_i is positive will be positive if $x > 1$.

To make a column positive in which a negative coefficient $-a_i$ occurs we must have

$$(a_0 + a_1)(x - 1) > a_2, \cdots, (a_0 + a_1 + \cdots + a_{r-1})(x - 1) > a_r,$$
etc.

Hence

$$x > \frac{a_2}{a_0 + a_1} + 1, \cdots, x > \frac{a_r}{a_0 + a_1 + \cdots + a_{r-1}} + 1, \text{ etc.}$$

Thus, to ensure every term being made positive, we must take the value of the greatest of the quantities found in this way. Such a value of x, therefore, is an upper bound of the positive roots.

8.3 Practical application. Theorem 3 will usually be found the more advantageous when the first negative coefficient is preceded by several positive coefficients, so that r is large. Theorem 4 may be found more advantageous when several large positive coefficients occur before the first large negative coefficient.

We use the least integer not below the numerical value given by either theorem as the upper bound.

Example. Find an upper bound of the roots of

$$x^7 + 5x^6 - 3x^5 + 3x^4 - 16x^3 - 9x^2 + 6x - 8 = 0.$$

Of the fractions

$$\frac{3}{1+5}, \quad \frac{16}{1+5+3}, \quad \frac{9}{1+5+3}, \quad \frac{8}{1+5+3+6},$$

the second is the greatest, and theorem 4 gives the integer 3 as an upper bound. Theorem 3 gives the integer 5 as an upper bound.

Exercises

Find an upper bound of the roots of the following equations:

1. $x^5 + 3x^4 + x^3 - 8x^2 - 51x + 18 = 0$
2. $x^8 + 19x^7 + 5x^6 - 15x^5 - 124x^4 + 9x^2 - 36 = 0$
3. $6x^5 - 3x^4 + 28x^3 + 61x^2 - 59x + 7 = 0$
4. $5x^4 - 21x^3 + 166x^2 - 546x + 306 = 0$
5. $6x^5 + 3x^4 - 96x^3 + 130x^2 + 20x - 32 = 0$
6. $3x^5 - 48x^3 + 50x^2 + 43x + 50 = 0$
7. $3x^6 + 2x^5 - 24x^3 - 20x^2 - 4 = 0$
8. $7x^6 + 18x^5 + 5x^4 - 3x^3 - 56x^2 - 27x - 3 = 0$
9. $8x^7 + 17x^5 + 5x^4 - 19x^3 - 128x^2 \quad 100x - 96 = 0$
10. $3x^6 - 12x^5 + 67x^4 + 10x^3 - 239x^2 + 7x - 48 = 0$

8.4 Newton's method. Theorem 5. *Any number which renders positive the polynomial $f(x)$ and all its derived functions $f'(x)$, $f''(x), \cdots, f^{(n)}(x)$ is an upper bound of the roots of $f(x) = 0$.*

This method is more laborious than the preceding methods but it has the advantage of always giving close upper bounds.

Proof: To prove the theorem, let the roots of $f(x) = 0$ be diminished by h; then $x - h = y$, and

$$f(x) = f(y + h) = f(h) + f'(h)y + \frac{f''(h)}{1\cdot 2}y^2 + \cdots + \frac{f^{(n)}(h)}{1\cdot 2\cdots n}y^n$$

If now h be such as to make all the coefficients

$$f(h), f'(h), f''(h), \cdots, f^{(n)}(h)$$

positive, the equation in y cannot have a positive root; that is, the equation in x has no root greater than h; hence h is an upper bound to the roots of $f(x) = 0$.

Example.

$$\begin{aligned} f(x) &= x^4 - 3x^3 - 4x^2 - 2x + 9 \\ f'(x) &= 4x^3 - 9x^2 - 8x - 2 \\ f''(x) &= 12x^2 - 18x - 8 \\ f'''(x) &= 24x - 18 \\ f^{IV}(x) &= 24. \end{aligned}$$

In this example $f'''(x)$ is positive for $x = 1$. This value of x makes $f''(x)$ negative. Increase x by one, $f''(x)$ is positive for $x = 2$. This value of x makes $f'(x)$ negative. Increase x by 1, $f'(x)$ is positive for $x = 3$. But this value of x makes $f(x)$ negative. Increase x by 1. We find that $x = 4$ makes $f(x)$ positive.

One advantage of Newton's method is that often, as in the present example, it furnishes two consecutive integers between which the greatest root lies. In the example given, the greatest root lies between 3 and 4.

The general method of procedure is as follows: Take the smallest integer that will make $f^{(n-1)}(x)$ positive. Proceeding upwards in order towards $f'(x)$, substitute this integer in the other functions of the series. When any function is reached which becomes negative for the integer in question, increase the integer successively by units, until an integer is found that makes that function positive. Proceed with the new integer as before. Continue in this way until an integer is obtained that makes all of the functions in the series positive.

It is assumed that when any number makes all of the derived functions up to a certain stage positive, any greater number will also make them positive.

This property is evident from the equation

$$g(a + h) = g(a) + g'(a)h + g''(a)\frac{h^2}{1\cdot 2} + \cdots$$

(taking $g(x)$ to represent any function in the series, and using the common notation for derived functions), which shows that if $g(a)$, $g'(a)$, $g''(a)$, $\cdots$ are all positive, and h is also positive, $g(a + h)$ must be positive.

8.5 Negative roots. To find bounds of the negative roots of an equation $f(x) = 0$ we put $-y$ for x, and then find bounds of the positive roots of the transformed equation in y; these bounds with their signs changed will be bounds of the negative roots of the given equation.

Example. $x^5 - 7x^4 - 18x^3 + 5x^2 + 4x + 36 = 0$. Put $x = -y$. We have

$$y^5 + 7y^4 - 18y^3 - 5y^2 + 4y - 36 = 0.$$

By theorem 4, we have $\dfrac{36}{1 + 7 + 4} + 1 = 4$ as an upper bound of the positive roots of the equation in y. By theorem 2, we have $\dfrac{36}{36 + 18} = \dfrac{2}{3}$ as a lower bound of the positive root of the equation in y. Thus the negative roots of the given equation must be between $-\frac{2}{3}$ and -4.

Exercises

1. Find by Newton's method upper bounds to the real roots of the following equations:

(a) $x^4 - x^3 - 5x^2 + 8x - 9 = 0$
(b) $x^4 - 5x^2 + 6x - 1 = 0$
(c) $x^4 - 2x^3 - 3x^2 - 15x - 3 = 0$
(d) $x^4 - 5x^2 - 6x - 1 = 0$
(e) $x^5 + x^4 - 4x^3 - 6x^2 - 700x + 500 = 0$
(f) $x^4 - 2x^3 - 3x^2 - 15x - 3 = 0$
(g) $5x^5 - 20x^4 - 10x^3 - 23x^2 - 90x - 297 = 0$
(h) $4x^5 - 15x^4 + 10x^3 - 32x^2 - 100x - 81 = 0$
(i) $3x^5 - 10x^4 - 10x^3 + 51x^2 - 36x - 24 = 0$
(j) $2x^4 - 6x^3 + 3x^2 - 60x - 30 = 0$

2. Show that the real roots of the following equations are between the bounds given:

(a) $x^4 - x^3 - 4x^2 - 3x + 1 = 0$; $-2, 3$
(b) $x^4 + x^3 - 10x^2 - x + 15 = 0$; $-4, 3$
(c) $x^5 + 5x^4 + x^3 - 16x^2 - 20x - 16 = 0$; $-5, 3$.
(d) $x^5 + x^4 + x^2 - 25x - 36 = 0$; $-3, 3$
(e) $x^5 + 5x^4 - 20x^2 - 19x - 2 = 0$; $-5, 3$

8.6 Theorem 6*. *If in the equation, with real coefficients,*

$$f(x) \equiv x^n + a_1x^{n-1} + \cdots + a_n = 0$$

where not all of the coefficients are positive, we strike out an arbitrary number of such pairs of terms as

$$a_px^{n-p} \text{ and } a_{p-q}x^{n-p+q}, \qquad a_p \geq |a_{p-q}|, \qquad a_{p-q} < 0,$$

and if, in the resulting equation, $a_{n-m}x^m$, where $a_{n-m} \neq 0$, is any term preceding all negative terms, $m - h$ is the power of the first negative term and g is the greatest of the numerical values of the negative coefficients, then each real root of $f(x) = 0$ is less than any term of the sequence

$$1 + \left(\frac{g}{a_{n-m}}\right)^{\frac{1}{h}} \equiv A_1; \quad \phi(A_1) \equiv A_2; \quad \phi(A_2) \equiv A_3, \cdots,$$

where

$$\phi(x) \equiv 1 + [(1 - x^{-m+h-1})g/a_{n-m}]^{\frac{1}{h}}, \qquad x > 1.$$

Proof. If $x > 1$, then

$$f(x) \geq a_{n-m}x^m - g(x^{m-h} + x^{m-h-1} + \cdots + 1)$$
$$= a_{n-m}x^m - g\frac{x^{m-h+1} - 1}{x - 1} \tag{2}$$

We seek a value of $x > 1$, such that

$$a_{n-m}x^m(x - 1) - g(x^{m-h+1} - 1) > 0. \tag{3}$$

Dividing this by x^{m-h+1} and transposing the g term gives

$$a_{n-m}x^{h-1}(x - 1) > g - gx^{-m+h-1},$$

whence

$$a_{n-m}(x - 1)^h > g - gx^{-m+h-1}$$

* Glenn James, *American Mathematical Monthly*, 1927, p. 351 ff.

which can be written

$$x > 1 + \left[\frac{g(1 - x^{-m+h-1})}{a_{n-m}}\right]^{\frac{1}{h}} \equiv \phi(x). \tag{4}$$

The right-hand member of the inequality in (4) defines $\phi(x)$. (4) is satisfied by $x = 1 + \left(\frac{g}{a_{n-m}}\right)^{\frac{1}{h}} \equiv A_1$, that is,

$$A_1 > \phi(A_1) \equiv A_2. \tag{5}$$

It follows from (4) and (5) that $\phi(A_1) > \phi(A_2) \equiv A_3$; similarly $\phi(A_2) > \phi(A_3)$, etc.

Consequently, if x is equal to or greater than any term of the sequence $A_1, A_2, A_3, \cdots$, then $f(x) > 0$, which proves the theorem.

Example 1. $x^3 - x^2 - x - 1 = 0$. We have $n = m = 3$, $h = g = 1$, $a_{n-m} = 1$. Hence the sequence of upper bounds is

$$1 + 1; 1 + (1 - 2^{-3}) = \frac{15}{8}; 2 - (8/15)^3 = 1.8483;$$

$$2 - (1.8483)^{-3} = 1.8416; 1.8399; 1.83944; 1.83932; 1.83931; \cdots.$$

The term 1.83931 is a root with four decimal places correct.

Example 2. $x^5 + 3x^4 - 2x^3 + 4x - 32 = 0$. Drop $3x^4 - 2x^3$. Then $n = m = 5$, $h = 5$, $g = 32$, $a_{n-m} = 1$. The sequence is

$$1 + 32^{\frac{1}{5}} = 3; 1 + 2(1 - 3^{-1})^{\frac{1}{5}} = 2.84$$

$$1 + 2[1 - (2.84)^{-1}]^{\frac{1}{5}} = 2.82.$$

Example 3. $x^5 - 2x^4 + 22x^3 + 60x^2 - 73x + 5 = 0$. This can be written $x^5 - 2x^4 + 22x^3 + 60x^2 - 51x - 22x + 5 = 0$. Drop the pairs $22x^3 - 22x$ and $60x^2 - 51x$. The first term of the sequence is 3.

Exercises

Find an upper bound to the real roots of the following equations:

1. $x^6 + 15x^5 - 16x^4 + 4x^3 + x - 8 = 0$
2. $x^5 - x^2 + 4x - 5 = 0$ (Drop $4x - 4$)
3. $x^4 + 5x^3 - 3x^2 + 2x - 16 = 0$ (Drop $5x^3 - 3x^2$)
4. $x^5 + 6x^4 - 5x^3 + 7x^2 - 4x - 32 = 0$ (Drop $6x^4 - 5x^3$ and $7x^2 - 4x$)

Chapter IX

SEPARATION OF THE ROOTS

9.1 Descartes' rule of signs. When two successive terms of an equation have different signs (ignoring all terms with zero coefficients), there is said to be a *variation of sign.* By the *number of variations of sign* of a rational integral equation is meant the total number of variations of sign presented by consecutive non-vanishing* terms.

Thus, in $x^6 + 3x^5 - 2x^4 - x^3 + 4x^2 + 5 = 0$, the second and third terms present one variation of sign. There is a second variation of sign presented by the fourth and fifth terms. The number of variations of sign of the equation is two.

Descartes' rule. *The number of positive real roots of a rational integral algebraic equation $f(x) = 0$, with real coefficients, is either equal to the number of the variations of sign in its coefficients or less than that number by a positive even integer.*

For example, the equation $x^6 + 3x^5 - 2x^4 - x^3 + 4x^2 + 5 = 0$ has two variations of sign and so by the rule has either two or no real positive roots. The equation $x^3 - x^2 + 2x - 3 = 0$ has three variations of sign, and so by the rule has either three or one real positive roots.

First, we shall show that if a polynomial $f(x)$ be multiplied by a factor $x - \alpha$ (α real), thereby introducing a new positive real root, the number of variations of sign of the product will exceed the number of variations of sign of the polynomial by an odd number.

Let the polynomial $f(x)$, which may be complete or incomplete, be arranged in descending powers of x. Let the coefficient of the term of highest degree be positive. For the moment assume that the coefficient of the term of lowest degree is positive. Then the signs of the terms vary in the following manner:

* An equation $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$, in which each of the $n + 1$ coefficients is different from zero, is said to be *complete*.

$$\begin{array}{cccccccccccc} & & 1 & & 2 & & 3 & & & v-1 & & v \\ + & \cdots & - & \cdots & + & \cdots & - & \cdots, & \text{etc.} & - & \cdots & + & \cdots, \end{array}$$

where the dots that follow a $+$ sign designate any number of consecutive terms whose coefficients are positive or zero, and the dots that follow a $-$ sign designate consecutive terms whose coefficients are negative or zero. The number 1 over the first $-$ sign shows that here occurs the first variation in sign. The number 2 above the second $+$ sign shows that here occurs the second variation in sign. The v over the last $+$ sign indicates that here occurs the v-th and last variation in sign.

Let α be a positive real root. Multiply $f(x)$ by $x - \alpha$, placing like powers of x in the same vertical column. We obtain a product whose signs may be written as follows:

$$\begin{array}{llllllll} ++\cdots & --\cdots & ++\cdots & --\cdots & \text{etc.} & --\cdots & ++\cdots & \\ \;\;-\cdots & -+\cdots & +-\cdots & -+\cdots & \text{etc.} & -+\cdots & +-\cdots & - \\ \hline +\pm\cdots & -\pm\cdots & +\pm\cdots & -\pm\cdots & \text{etc.} & -\pm\cdots & +\pm\cdots & - \\ & 1 & 2 & 3 & & v-1 & v & v+1 \\ A_0 & A_1 & A_2 & A_3 & & A_{v-1} & A_v & A_{v+1} \end{array}$$

The $\pm$ indicates that the sign of that term is undetermined. We say that it is ambiguous; it may be positive, negative, or zero. Each permanence of sign gives rise to an ambiguity: the dots following $\pm$ indicate ambiguities. It is clear, however, that to each variation of sign in $f(x)$ there corresponds a variation of sign in $(x - \alpha)\cdot f(x)$. These variations are indicated by the numbers 1, 2, 3, $v - 1$, v, beneath the signs in the product. Further, in the product, there is introduced an additional variation of sign at the end. Hence the product contains at least one more variation of sign than the polynomial $f(x)$.

Any sequence of signs beginning with a $+$ and ending with a $-$ has an odd number of variations of sign. These sequences are such as the one from the $+$ sign above A_0 to the $-$ sign above A_1, and the sequence from the $+$ above A_2 to the $-$ above A_3. Any sequence of signs beginning with a $-$ and ending with a $+$ has an odd number of variations of sign. Such sequences are from the $-$ above A_1 to the $+$ above A_2 and from the $-$ above A_{v-1} to the $+$ above A_v. There are $v + 1$ such sequences. In each sequence there are an odd number of changes of sign; that is,

$1 + 2k$ changes of sign. Then the total number of variations of sign in $(x - \alpha)\cdot f(x)$ is

$$(1 + 2k_1) + (1 + 2k_2) + \cdots + (1 + 2k_{v+1}) = v + 1 + \text{an even number.}$$

That is, the number of variations of sign in $(x - \alpha)\cdot f(x)$ is either one more than the number of variations of sign in $f(x)$ or exceeds the number of variations of sign in $f(x)$ by one plus an even positive integer.

The same conclusion is reached when the last term in $f(x)$ is assumed to be negative.

We now use the above to prove Descartes' rule.

Let $g(x)$ be the product of all those factors corresponding to negative and complex roots of $f(x) = 0$. Since $g(x) = 0$ has no positive roots, the first and last terms in $g(x)$ have like signs, which may be taken as positive. Then the number of variations of sign in $g(x)$ is an even number $2h$ where h is either zero or a positive integer. Now, if $g(x)$ be multiplied by $x - \alpha_1$, where α_1 is a positive root, we obtain in the product $2h + (1 + 2h_1)$ variations of sign. In the same way a second factor $x - \alpha_2$ gives rise to $1 + 2h_2$ additional variations of sign, so that in the product $(x - \alpha_1)(x - \alpha_2)g(x)$ there are $2h + (1 + 2h_1) + (1 + 2h_2)$ variations of sign. Continuing in this way, after the introduction of r positive roots the product

$$(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_r)g(x) = f(x)$$

will have

$$2h + (1 + 2h_1) + (1 + 2h_2) + \cdots + (1 + 2h_r)$$

variations of sign; that is, the number of variations of sign is r plus a positive even integer. This completes the proof of the theorem.

To make use of Descartes' rule to determine a superior bound to the negative roots of $f(x) = 0$, we set up an equation whose roots are the roots of $f(x) = 0$ *with their signs changed.* The new equation can be obtained from $f(x) = 0$ by replacing x by $-x$; that is, we apply Descartes' rule to the equation $f(-x) = 0$.

Corollary: The number of negative roots of an algebraic equation $f(x) = 0$, with real coefficients, is either equal to the number of variations of sign of $f(-x) = 0$ or less than that number by a positive even integer.

Exercises

1. Determine the nature of the roots of $x^3 + 5x - 3 = 0$.

There is one variation of sign; therefore one positive root. Replace x by $-x$. We have $x^3 + 5x + 3 = 0$. This new equation has no variation of sign; therefore no positive root. Hence the original equation has no negative root. The given equation has, then, only one real root which is positive; the other two roots must be imaginary, and in particular will be mutually conjugate.

2. Determine the nature of the roots of $f(x) = x^4 + 7x^2 + 3x - 5 = 0$.

$f(x)$ has one variation of sign and $f(-x)$ has one variation of sign. Hence $f(x) = 0$ has one positive root, one negative root, and two imaginary roots.

Prove by Descartes' rule the following statements:

3. $x^3 + 2x + 3 = 0$ has two imaginary roots.
4. $x^3 + px + q = 0$ $(p, q > 0)$ has two imaginary roots.
5. $x^3 + px - q = 0$ $(p, q > 0)$ has two imaginary roots.
6. $x^4 + x^2 + 2x - 3 = 0$ has one positive root, one negative root, two imaginary roots.
7. $x^3 + x^2 + x + 1 = 0$ has two imaginary roots.
   Hint: multiply by $x - 1$.
8. $x^6 + 3x^4 + 5x^2 + 1 = 0$ has six imaginary roots.
9. $x^5 + x^3 + 5x = 0$ has four imaginary roots. The only real root is $x = 0$.
10. $x^4 - 2x^3 + 3x^2 - 4x + 1 = 0$ has no negative root.
11. $x^6 - 3x^4 - 4x + 5 = 0$ has at least two imaginary roots.
12. $x^5 + 3x^2 + 1 = 0$ has four imaginary roots.
13. $3x^6 + 5x^4 + 3x^2 + 4 = 0$ has six imaginary roots.
14. $x^6 - 7 = 0$ has one positive and one negative root. Four imaginary roots.
15. $3x^5 - 8 = 0$ has one positive root and four imaginary roots.
16. $x^4 + 3x^3 + 2x - 1 = 0$ has one positive and one negative root, and two imaginary roots.
17. $x^4 + 3x^2 - 2x + 1 = 0$ has not more than two positive roots; at least two imaginary roots and no negative roots.
18. $x^4 - 3x^2 + 2x + 1 = 0$ has not more than two positive roots and not more than two negative roots.
19. An equation in which all of the coefficients have the same sign has no positive root.
20. For n even, $x^n - 1 = 0$ has $n - 2$ imaginary roots.

21. For n odd, $x^n - 1 = 0$ has $n - 1$ imaginary roots.
22. For n even, $x^n + 1 = 0$ has no real roots.
23. For n odd, $x^n + 1 = 0$ has $n - 1$ imaginary roots.
24. If the signs of the terms of any complete equation are alternately positive and negative, it cannot have a negative root.
25. If an equation contains a series of terms all of which have positive coefficients followed by a series all of which have negative coefficients, it has only one positive root.
26. If an equation contains only even powers of x, and if all of the coefficients are positive, it cannot have a real root.
27. If a rational integral algebraic equation contains only odd powers of x, and if all of the coefficients are positive, it has the root zero and no other real root.
28. If a rational integral algebraic equation $f(x) = 0$ is complete and all of its roots are real, then the number of positive roots is equal to the number of variations of sign in $f(x) = 0$ and the number of negative roots is equal to the number of variations of sign in $f(-x) = 0$.
29. In astronomy,** the problem of three bodies gives rise to the equation

$$r^5 + (3 - \mu)r^4 + (3 - 2\mu)r^3 - \mu r^2 - 2\mu r - \mu = 0, \quad (0 < \mu < 1).$$

Use Descartes' rule to show that this equation has a single positive real root.

9.2 Budan's theorem. *Let $f(x) = 0$ be a rational integral algebraic equation of degree n, with real coefficients. Let two real numbers a and b, $a < b$, neither a root of $f(x) = 0$, be substituted in the series formed by $f(x)$ and its successive derived functions, viz.,*

$$f(x), f'(x), f''(x), \quad , f^{(n)}(x); \tag{1}$$

then the excess of the number of variations of sign in the series (1) when $x = a$, over the number of variations of sign when $x = b$, either equals the number of real roots of $f(x) = 0$ between a and b or exceeds the number of roots by a positive even integer. A root of multiplicity m is here counted as m roots.

We shall call the whole series (1) Budan's functions.

** F. R. Moulton, *Celestial Mechanics*, p. 197, The Macmillan Company, 1902.

It is convenient at times to write Budan's functions with subscripts to indicate derivatives:

$$f(x), f_1(x), f_2(x), \cdots, f_n(x).$$

These should not be confused with Sturm's functions (9.3).

No change can occur in the sign of any one of Budan's functions except when x passes through a value which makes that function vanish. The following four cases can arise:

Case I: The value of x may pass through a single root of $f(x) = 0$. Budan's functions lose one change of sign; for $f(x)$ and $f'(x)$ have unlike signs immediately before, and like signs immediately after, the passage through a root of $f(x) = 0$, while the functions $f_2(x), \cdots, f_n(x)$ do not change their signs, by hypothesis.

Case II: The value of x may pass through a root c occurring r times in $f(x) = 0$. That is,

$$f(c) = f_1(c) = \cdots = f_{r-1}(c) = 0; \qquad f_r(c) \neq 0.$$

From Taylor's theorem, equation (3), Chapter IV, we have

$$\begin{aligned}
f(c-h) &= f(c) + f_1(c)(-h) + \tfrac{1}{2}f_2(c)(-h)^2 + \cdots \\
&\qquad + \frac{f_r(c)}{\underline{|r}}(-h)^r + \cdots \\
f_1(c-h) &= f_1(c) + f_2(c)(-h) + \cdots + \frac{f_r(c)}{\underline{|r-1}}(-h)^{r-1} + \cdots \\
f_2(c-h) &= f_2(c) + \cdots + \frac{f_r(c)}{\underline{|r-2}}(-h)^{r-2} + \cdots \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
f_{r-1}(c-h) &= \qquad f_{r-1}(c) + f_r(c)(-h) + \cdots \\
f_r(c-h) &= \qquad f_r(c) + \cdots.
\end{aligned} \tag{2}$$

We may take h so small that the signs of the series of terms $f(c-h), f_1(c-h), f_2(c-h), \cdots, f_{r-1}(c-h), f_r(c-h)$ shall be respectively the same as the signs of the series

$$f_r(c)(-h)^r, f_r(c)(-h)^{r-1}, f_r(c)(-h)^{r-2}, \cdots, f_r(c)(-h), f_r(c).$$

The signs of this series alternate.

Rewrite equations (2) with $-h$ replaced by $+h$. Then we see that for $x = c - h$, the first $r + 1$ of Budan's functions have signs

alternately + and −, and that for $x = c + h$ these same $r + 1$ functions all have the same sign, namely that of $f_r(x)$. Hence the number of changes of sign that are lost is r, which is the degree of multiplicity of the root.

Case III: Suppose when $x = c$ that one of the derived functions vanishes, but neither of the two adjacent functions; that is,

$$f_r(c) = 0, \qquad f_{r-1}(c) \neq 0, \qquad f_{r+1}(c) \neq 0.$$

Then if h is taken small enough, when $x = c - h$, the signs of the three terms $f_{r-1}(x)$, $f_r(x)$, $f_{r+1}(x)$, are respectively the same as the signs of $f_{r-1}(c)$, $-hf_{r+1}(c)$, $f_{r+1}(c)$, and when $x = c + h$, the signs are the same as the signs of $f_{r-1}(c)$, $hf_{r+1}(c)$, $f_{r+1}(c)$. Thus if $f_{r-1}(c)$ and $f_{r+1}(c)$ have the same sign, Budan's functions lose *two* changes of sign as x increases through c, and if $f_{r-1}(c)$ and $f_{r+1}(c)$ have contrary signs, Budan's functions neither gain nor lose a change of sign.

That is, if any changes of sign are lost it is an even number two.

Case IV: Suppose when $x = c$ that m successive derived functions vanish; for example

$$f_r(c) = f_{r+1}(c) = f_{r+2}(c) = \cdots = f_{r+m-1}(c) = 0$$

while $f_{r-1}(c) \neq 0$ and $f_{r+m}(c) \neq 0$.

By Taylor's theorem we have

$$\begin{aligned} f_{r-1}(c-h) &= f_{r-1}(c) + \cdots \\ f_r(c-h) &= f_{r+m}(c)(-h)^m + \cdots \\ f_{r+1}(c-h) &= f_{r+m}(c)(-h)^{m-1} + \cdots \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ f_{r+m-1}(c-h) &= f_{r+m}(c)(-h) + \cdots \\ f_{r+m}(c-h) &= f_{r+m}(c) + \cdots . \end{aligned} \tag{3}$$

If both $f_{r-1}(c)$ and $f_{r+m}(c)$ are positive, the following table, which follows from equations (3), represents the situation. If both are negative, all signs in the table are to be changed.

f_{r-1}	f_r	f_{r+1}	f_{r+2}	$\cdots$	f_{r+m-3}	f_{r+m-2}	f_{r+m-1}	f_{r+m}
+	$(-)^m$	$(-)^{m-1}$	$(-)^{m-2}$	$\cdots$	−	+	−	+.

If f_{r+m} is negative and f_{r-1} is positive, the following table, which follows from (3), represents the situation. If f_{r+m} is positive and f_{r-1} is negative, all signs in the table are to be changed.

$$\begin{array}{cccccccc} f_{r-1} & f_r & f_{r+1} & f_{r+2} \cdots & f_{r+m-3} & f_{r+m-2} & f_{r+m-1} & f_{r+m} \\ + & (-)^{m+1} & (-)^m & (-)^{m-1} \cdots & + & - & + & -. \end{array}$$

By taking h small enough and positive, we obtain the following results with respect to the $m + 2$ terms $f_{r-1}, f_r, \cdots, f_{r+m}$.

(a) Let m be *even*. If $f_{r-1}(c)$ and $f_{r+m}(c)$ have the *same* sign, the terms present m changes of sign when $x = c - h$, and no changes of sign when $x = c + h$. If $f_{r-1}(c)$ and $f_{r+m}(c)$ have *unlike* signs, the terms present $m + 1$ changes of sign when $x = c - h$, and one change of sign when $x = c + h$. Thus in both cases Budan's functions lose m (an even number) changes of sign as x increases through c.

(b) Let m be *odd*. If $f_{r-1}(c)$ and $f_{r+m}(c)$ have the *same* sign, the terms present $m + 1$ changes of sign when $x = c - h$, and no change of sign when $x = c + h$. Thus Budan's functions lose $m + 1$ (an even number) changes of sign as x increases through c. If $f_{r-1}(c)$ and $f_{r+m}(c)$ have *unlike* signs, the terms present m changes of sign when $x = c - h$, and one change of sign when $x = c + h$. Thus Budan's functions lose $m - 1$ (an even number) changes of sign as x increases through c.

We conclude that as x increases from a to b no change of sign can be gained; that for each passage through a single root of $f(x) = 0$ one change is lost; and that under no circumstances except through a root of $f(x) = 0$ can an odd number of changes be lost. Hence the number of changes lost during the whole change of x from a to b must be either equal to the number of real roots of $f(x) = 0$ in the interval, or must exceed it by an even number.

Budan's method has one advantage over that of Sturm (9.3), in that Budan's functions are easily obtained. Sturm's method is superior in that it always gives the exact number of real roots between a and b.

Descartes' rule of signs is a corollary to Budan's theorem. This can be seen as follows: Let

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0,$$

where the coefficients are real and $a_0 > 0$. For $x = 0$, Budan's functions $f, f_1, f_2, \cdots, f_{n-1}, f_n$ and the coefficients $a_n, a_{n-1}, a_{n-2}, \cdots, a_1, a_0$ have the same set of signs. For $x = +\infty$ the functions are all positive. Then the excess of the number of variations of sign of the functions for $x = 0$ over the number of variations of sign for $x = \infty$ is the same as the number of variations of sign of $f(x)$.

Example. Locate the roots of

$$\begin{aligned} f(x) &= x^5 + x^4 - 4x^3 - 3x^2 + 3x + 1 = 0. \quad \text{We have} \\ f_1(x) &= 5x^4 + 4x^3 - 12x^2 - 6x + 3 \\ f_2(x) &= 20x^3 + 12x^2 - 24x - 6 \\ f_3(x) &= 60x^2 + 24x - 24 \\ f_4(x) &= 120x + 24 \\ f_5(x) &= 120 \end{aligned}$$

We form the following table:

x	f	f_1	f_2	f_3	f_4	f_5
-2	$-$	$+$	$-$	$+$	$-$	$+$
-1	$-$	$-$	$+$	$+$	$-$	$+$
0	$+$	$+$	$-$	$-$	$+$	$+$
1	$-$	$-$	$+$	$+$	$+$	$+$
2	$+$	$+$	$+$	$+$	$+$	$+$

From this table we find that there may be two roots in the interval $(-2, -1)$ and must be one root in each of the intervals $(-1, 0)$; $(0, 1)$; $(1, 2)$.

Exercises

Verify the following statements:

1. The real roots of $x^3 + x^2 - 2x - 1 = 0$ are in the intervals $(-2, -1)$; $(-1, 0)$; $(1, 2)$.

2. The real roots of $x^5 + x^4 + x^2 - 25x - 36 = 0$ are in the intervals $(-3, -2)$; $(-2, -1)$; $(2, 3)$.

3. The equation $x^4 - 2x^3 - 7x^2 + 10x + 10 = 0$ has one real root in each of the intervals $(-3, -2)$; $(-1, 0)$; and two roots between 2 and 3.

4. The real roots of $3x^4 - 6x^2 - 8x - 3 = 0$ are in the intervals $(-1, 0)$; $(1,2)$.

5. The real roots of $x^4 - 16x^3 + 69x^2 - 70x - 42 = 0$ are in the intervals $(-1, 0)$; $(2, 3)$; $(4,5)$; $(9, 10)$.

6. The real roots of $x^4 - 5x^3 + 10x^2 - 6x - 21 = 0$ are in the intervals $(-1, 0)$; $(3, 4)$.

7. The equation $x^5 - 10x^3 + 6x + 1 = 0$ has single roots in the intervals $(-4,-3)$; $(0, 1)$; $(3, 4)$ and two roots in the interval $(-1, 0)$.

8. The equation $x^5 - 5x^4 + 5x^3 + 5x^2 - 5x - 1 = 0$ has one real root in the interval $(0, 2)$ namely $x = 1$; and two real roots in each of the intervals $(-1, 0)$; $(2, 3)$.

9.3 Sturm's functions. Let $f(x) = 0$ be a rational integral algebraic equation which has no equal roots. Form the first derived function namely $f'(x)$. Now proceed in the usual manner with the process of finding the greatest common divisor of $f(x)$ and $f'(x)$, with the exceptions noted below. In the first step divide $f(x)$ by $f'(x)$ until you obtain a remainder $r(x)$ whose degree is less than that of $f'(x)$. If $q_1(x)$ is the quotient, then $f(x) = q_1(x)f'(x) + r(x)$. Instead of dividing $f'(x)$ by $r(x)$, divide $f'(x)$ by $f_2(x) = -r(x)$. At each step *the sign of the remainder must be changed* before it is used as a divisor. Continue this process until a remainder is obtained that does not contain x and change the sign of this last remainder also. This process is just as effective as the usual procedure in finding the greatest common divisor. Designate the remainders with their signs changed by $f_2(x), f_3(x), \cdots, f_n(x)$. Let $q_1, q_2, \cdots, q_{n-1}$ be the successive quotients in the process, then

$$\begin{aligned}
f(x) &= q_1 f'(x) - f_2(x) \\
f'(x) &= q_2 f_2(x) - f_3(x) \\
f_2(x) &= q_3 f_3(x) - f_4(x) \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
f_{n-2}(x) &= q_{n-1} f_{n-1}(x) - f_n(x).
\end{aligned}$$

The last remainder $f_n(x)$ is a non-zero constant.

The functions $f(x), f'(x), f_2(x), f_3(x), \cdots, f_n(x)$ are called *Sturm's functions*. $f_0(x) \equiv f(x); f_1(x) \equiv f'(x)$.

9.4 Sturm's theorem. *Let $f(x) = 0$ be a rational integral algebraic equation, with real coefficients and without multiple roots.*

Let any two real numbers a and b, $a < b$, neither a root of $f(x) = 0$, be substituted for x in the series of $n + 1$ functions

$$f(x), f'(x), f_2(x), f_3(x), \cdots, f_n(x), \tag{4}$$

consisting of the given equation $f(x)$, its first derived function $f'(x)$, and the successive remainders, with their signs changed, in the process of finding the greatest common divisor of $f(x)$ and $f'(x)$. Then the exact number of real roots of $f(x) = 0$ between $x = a$ and $x = b$ is exactly equal to the number of variations of sign in the series when $x = a$, diminished by the number of variations of sign in the series when $x = b$.

In the first place, let us observe that in the case under consideration no two consecutive functions in the series can vanish for the same value of x; that is, have a common factor. For, if any two consecutive functions of the series had a common factor, then $f(x)$ and $f'(x)$ would have this same common factor and hence $f(x)$ would have a multiple root, which is contrary to our hypothesis.

In the second place, we observe that as x moves from a to b, the number of variations of sign of the series of functions (4) for $x = \alpha$ is the same as the number of variations of sign for $x = \beta$, where α and β are real numbers between a and b such that no one of the functions (4) vanishes between α and β or for $x = \alpha$ or for $x = \beta$. For, if any one of the functions (1) is positive (negative) for $x = \alpha$ and does not vanish between $x = \alpha$ and $x = \beta$ or for $x = \beta$, then for $x = \beta$ the function is still positive (negative).

Hence, during the passage of x from a to b, the only cases in which there can be any changes in the number of variations of sign of the series of functions (4) are the following:

Case I: When x passes through a value which causes one of the functions $f', f_2, \cdots, f_{n-1}$ to vanish.

Case II: When x passes through a root of $f(x) = 0$.

Case I: Suppose x, in passing from $x = a$ to $x = b$, takes a value α such that $f_r(\alpha) = 0$. From the equation

$$f_{r-1}(x) = q_r f_r(x) - f_{r+1}(x)$$

we have

$$f_{r-1}(\alpha) = -f_{r+1}(\alpha),$$

which shows that $x = \alpha$ gives to $f_{r-1}(x)$ and $f_{r+1}(x)$ the same numerical value with different signs. In passing from a value a little

less than α to a value a little greater than α we can consider the interval so small that it contains no root of $f_{r-1}(x)$ or $f_{r+1}(x)$; hence, throughout the interval, these two functions do not change sign. If the sign of $f_r(x)$ does not change in this interval, there is no change in the series of signs. If the sign of $f_r(x)$ does change in this interval, no variation of sign is gained or lost in the group of three. For, no matter which sign is placed between two unlike signs, there is only one variation of sign. The reader may check for himself. The only possibilities are a change from

$$\begin{array}{ccccccc} + & + & - & \text{to} & + & - & - \\ + & - & - & \text{to} & + & + & - \\ - & + & + & \text{to} & - & - & + \\ - & - & + & \text{to} & - & + & + \end{array}$$

Hence, no variation is either gained or lost among Sturm's functions.

Case II: When x, in passing from a to b, takes a value r which is a root of $f(x) = 0$, then Sturm's functions lose *one* variation of sign. The proof is as follows:

By Taylor's theorem, §4.3,

$$f(r - h) - f(r) = -hf'(r) + \tfrac{1}{2}h^2f''(r) - \cdots,$$

$$f(r + h) - f(r) = hf'(r) + \tfrac{1}{2}h^2f''(r) + \cdots.$$

The right-hand member of both equations is a polynomial in h. For sufficiently small values of h each polynomial in h will have the same sign as its first term. Hence, when $f(r) = 0$, if $f'(r)$ is positive, then $f(r - h)$ is negative, and $f(r + h)$ is positive. That is, the signs of $f(x)$ and $f'(x)$ will be $-+$ just before $x = r$, and $++$ just after $x = r$, and one variation of sign is lost. If $f'(r)$ is negative, then $f(r - h)$ is positive and $f(r + h)$ is negative. That is, the signs of $f(x)$ and $f'(x)$ will be $+-$ just before $x = r$, and $--$ just after $x = r$, and one variation of sign is lost. Hence, whether $f'(r)$ is positive or negative, the sequence $f(x), f'(x)$ shows one loss in variation of sign as x passes from $r - h$ through r to $r + h$.

We have now shown that as x increases and passes from a to b, there are no changes in the number of variations of sign of the series of functions (4), with the exception that one variation is *lost* every time that x passes through a value which causes $f(x)$ to vanish. Hence, the number of variations lost as x increases

from a to b is equal to the number of real roots of $f(x) = 0$ between a and b.

9.5 Sturm's theorem for multiple roots. *If a and b are real numbers, $a < b$, neither a root of $f(x) = 0$, and Sturm's functions are represented by the series $f(x), f'(x), f_2(x), \cdots, f_n(x)$, where $f_n(x)$ is the highest common factor of $f(x)$ and $f'(x)$, then the excess of the number of variations of sign in this series when $x = a$, over the number of variations of sign when $x = b$, is equal to the number of real roots between a and b, each multiple root counting only once.*

In the case of equal roots $f(x)$ and $f'(x)$ have a common factor; hence, $f_n(x)$, the last of Sturm's functions, is not a constant, but the greatest common divisor of $f(x)$ and $f'(x)$.

Suppose α is a multiple root of $f(x) = 0$ of order m, then

$$\begin{aligned} f(x) &= (x - \alpha)^m (x - r_1)(x - r_2) \cdots \\ f'(x) &= m(x - \alpha)^{m-1}(x - r_1)(x - r_2) \cdots \\ &\quad + (x - \alpha)^m (x - r_2)(x - r_3) \cdots \\ &\quad + (x - \alpha)^m (x - r_1)(x - r_3) \cdots \\ &\quad + \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned}$$

From these equations we see that $f(x)$ and $f'(x)$ have the common factor $(x - \alpha)^{m-1}$.

Let us consider the equations

$$\begin{aligned} f(x) &= q_1 f'(x) - f_2(x) \\ f'(x) &= q_2 f_2(x) - f_3(x) \\ f_2(x) &= q_3 f_3(x) - f_4(x) \\ & \\ f_{n-2}(x) &= q_{n-1} f_{n-1}(x) - f_n(x). \end{aligned} \tag{5}$$

Since $f(x)$ and $f'(x)$ have the common factor $(x - \alpha)^{m-1}$, we see from equations (5) that $(x - \alpha)^{m-1}$ is also a common factor of each of the functions $f(x), f'(x), f_2(x), f_3(x), \cdots, f_n(x)$. Divide each of equations (5) by $(x - \alpha)^{m-1}$. We obtain a new set of equations which may be represented as follows:

$$\begin{aligned} \phi(x) &= q_1 \phi_1(x) - \phi_2(x) \\ \phi_1(x) &= q_2 \phi_2(x) - \phi_3(x) \\ \phi_2(x) &= q_3 \phi_3(x) - \phi_4(x) \\ \cdot \quad \cdot \quad & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \phi_{n-2}(x) &= q_{n-1} \phi_{n-1}(x) - \phi_n(x). \end{aligned} \tag{6}$$

Note that $\phi_1(x) = m\phi'(x)$, then for any value of x, $\phi_1(x)$ and $\phi'(x)$ have the same sign.

We may use $\phi(x)$, $\phi_1(x)$, $\cdots$, $\phi_n(x)$ as Sturm's functions to determine the exact number of simple real roots of $\phi(x) = 0$ between a and b. This will give the exact number of roots of $f(x) = 0$ between a and b, each multiple root being counted only once.

The number of variations of sign will always be the same for the series

$$f(x),\ f'(x),\ f_2(x),\ \cdots,\ f_n(x)$$

as for

$$\phi(x),\ \phi_1(x),\ \phi_2(x),\ \cdots,\ \phi_n(x).$$

For, corresponding terms of the two series differ only by the factor $(x - \alpha)^{m-1}$. Hence, for any value of x, the signs of the terms in the first series are all the same as those of the second series, or the signs are all different.

By finding the number of variations of sign in the second series we can determine the exact number of simple real roots of the equation $\phi(x) = 0$ between a and b, and this number is the exact number of real and distinct roots of $f(x) = 0$ between a and b.

If there is more than one multiple root, by a repetition of the process of deriving equations (6), one can arrive at a set like (6) in which $\phi(x)$ has only simple roots and there is present in $\phi(x)$ every real and distinct root of $f(x) = 0$.

Descartes' rule gives only a superior bound for the number of positive and negative roots. Sturm's method gives always the *exact* number of real and distinct roots between a and b.

In order to avoid fractions, we may multiply $f(x)$ by a positive constant before dividing by $f'(x)$, and multiply any $f_i(x)$ by a positive constant before dividing by $f_{i+1}(x)$. Also before using any $f_i(x)$ as a divisor we may remove from it any constant positive factor.

Exercises

1. Find the situation of the real roots of $x^3 - 7x + 5 = 0$. We find

$$f'(x) = 3x^2 - 7; \qquad f_2(x) = 14x - 15; \qquad f_3(x) = 697.$$

We give the signs of Sturm's functions for the indicated values of x:

x	$f(x)$	$f'(x)$	$f_2(x)$	$f_3(x)$
-3	$-$	$+$	$-$	$+$
-2	$+$	$+$	$-$	$+$
-1	$+$	$-$	$-$	$+$
0	$+$	$-$	$-$	$+$
1	$-$	$-$	$-$	$+$
2	$-$	$+$	$+$	$+$
3	$+$	$+$	$+$	$+$

There are three variations of sign in Sturm's functions for $x = -3$ and two variations for $x = -2$; hence there is one root between -3 and -2. There are two variations of sign for $x = 0$ and one variation for $x = 1$; hence there is one root between 0 and 1. There is one variation of sign for $x = 2$ and none for $x = 3$; hence there is one root between 2 and 3.

$f(x)$ is $-$ for $x = 2$ and $+$ for $x = 3$; hence there is an odd number of roots between 2 and 3. Sturm's functions show that there is exactly one root between 2 and 3.

2. Locate the real roots of $x^4 - 2x^3 - 3x^2 + 10x - 4 = 0$. We find

$$f'(x) = 4x^3 - 6x^2 - 6x + 10; \qquad f_2(x) = 9x^2 - 27x + 11;$$
$$f_3(x) = -8x - 3; \qquad f_4(x) = -1433.$$

We have the following series of signs for Sturm's functions:

$-\infty$	$+$	$-$	$+$	$+$	$-$,
0	$-$	$+$	$+$	$-$	$-$,
∞	$+$	$+$	$+$	$-$	$-$.

Accordingly, there are two real roots, one positive, and one negative, and two imaginary roots. To find the position of the real roots, it is sufficient to substitute positive and negative integers successively in $f(x)$ alone, since there is only one positive and one negative root. We find that the negative root is between -2 and -3, and the positive root between 0 and 1.

By means of Sturm's functions verify that the real roots of the following equations are in the interval indicated.

3. $x^3 - 7x + 7 = 0$. $(-4, -3)$. Two in the interval $(1, 2)$.
4. $x^3 - 7x + 14 = 0$. $(-4, -3)$. Two imaginary roots.

5. $x^4 - 4x^3 + 6x^2 - 8x + 1 = 0$. (0, 1), (2, 3). Two imaginary roots.

6. $x^4 - 12x^2 + 12x - 3 = 0$. (−4, −3), (2, 3). Two in the interval (0, 1).

7. $8x^3 - 12x^2 + 1 = 0$. (−1, 0), (0, 1), (1, 2). The roots are −0.27; 0.33; 1.44.

8. $16x^4 - 16x^2 + 3 = 0$. (−1, −0.6), (−0.6, 0), (0, 0.6), (0.6, 1). The roots are ±0.5, ±0.87.

9. $x^6 - 1.771561 = 0$. (−2, −1), (1, 2). Four imaginary roots. The real roots are ±1.1.

10. $x^6 - 3x^4 + 2.620739 = 0$. Two roots between 1 and 2. Two roots between −1 and −2. Two imaginary roots. Two of the real roots are ±1.1.

11. Show that for the cubic $x^3 + px + q = 0$, we have

$$f'(x) = 3x^2 + p; \quad f_2(x) = -2px - 3q; \quad f_3(x) = -4p^3 - 27q^2.$$

Hint: Multiply $f'(x)$ by $4p^2$ before dividing by $f_2(x)$.

12. Show that for the quartic $x^4 + px^2 + qx + r = 0$, we have

$$f'(x) = 4x^3 + 2px + q; \qquad f_2(x) = -2px^2 - 3qx - 4r;$$

$$f_3(x) = (8pr - 2p^3 - 9q^2)x - 12qr - qp^2.$$

Hint: Divide $p^2f'(x)$ by $f_2(x)$. To obtain $f_4(x)$, divide

$$(8pr - 2p^3 - 9q^2)^2 f_2(x) \text{ by } f_3(x).$$

13. Locate the real roots of $4x^4 - 12x^3 + 13x^2 - 12x + 9 = 0$. We find

$$f'(x) = 16x^3 - 36x^2 + 26x - 12; \qquad f_2(x) = 2x^2 + 33x - 54;$$

$$f_3(x) = -2x + 3.$$

Sturm's functions stop with $f_3(x)$, thus establishing the existence of equal roots. $2x - 3$ is a common factor of $f(x)$ and $f'(x)$ and hence $x = \frac{3}{2}$ is a double root of $f(x) = 0$. For Sturm's functions we have

x	$f(x)$	$f'(x)$	$f_2(x)$	$f_3(x)$
$-\infty$	+	−	+	+
$+\infty$	+	+	+	−

The equation has only one real distinct root; the other two roots are imaginary. Factoring, we find $f(x) \equiv (2x - 3)^2(x^2 + 1)$.

Verify the following:

14. For $f(x) = x^4 - 5x^3 + 9x^2 - 7x + 2 = 0$, Sturm's functions are

$$f'(x) = 4x^3 - 15x^2 + 18x - 7,$$

$$f_2(x) = x^2 - 2x + 1.$$

$x = 1$ is a triple root. The other root is $x = 2$.

15. For $f(x) = x^4 - 6x^3 + 13x^2 - 12x + 4 = 0$, Sturm's functions are

$$f'(x) = 4x^3 - 18x^2 + 26x - 12$$

$$f_2(x) = x^2 - 3x + 2 = (x - 1)(x - 2).$$

Each of the roots 1, 2 is a double root.

16. For $f(x) = x^6 - 7x^5 + 15x^4 - 40x^2 + 48x - 16 = 0$, Sturm's functions are

$$f'(x) = 6x^5 - 35x^4 + 60x^3 - 80x + 48,$$

$$f_2(x) = 13x^4 - 84x^3 + 192x^2 - 176x + 48,$$

$$f_3(x) = x^3 - 6x^2 + 12x - 8 = (x - 2)^3.$$

$x = 2$ is a quadruple root. There is one real root in the interval $(-2, -1)$ and one real root in the interval $(0, 1)$.

17. For $f(x) = 32x^5 - 16x^4 - 16x^3 + 8x^2 + 2x - 1 = 0$, Sturm's functions are

$$f'(x) = 160x^4 - 64x^3 - 48x^2 + 16x + 2$$

$$f_2(x) = 8x^3 - 4x^2 - 2x + 1 = (2x - 1)^2(2x + 1).$$

Hence, $x = \frac{1}{2}$ is a triple root and $x = -\frac{1}{2}$ is a double root.

Chapter X

SOLUTION OF NUMERICAL EQUATIONS

10.1 Horner's method. The first step in finding the numerical value of a real root of a rational integral algebraic equation is to isolate the root. Then, whether the root is rational or irrational, and expressible as a terminating or nonterminating decimal, it can be obtained by Horner's method. The root is obtained figure by figure; first the integral part (if any), and then the decimal part, till the work terminates if the root is a terminating decimal, or to any desired number of decimal places otherwise.

The main principle in Horner's method is the successive diminution of the roots of the given equation by known amounts using synthetic division (*see* §4.7).

The method will be illustrated by the following example:

Let us find the root between 2 and 3 of

$$f(x) = x^3 + x^2 - 6x - 1 = 0.$$

The first step even if the root were to lie very near to 3, is to diminish the roots by 2 as follows:

$$\begin{array}{rrrr|l}
1 & +1 & -6 & -1 & \underline{2} \\
 & +2 & +6 & 0 & \\
\hline
1 & +3 & 0 & -1 & \\
 & +2 & +10 & & \\
\hline
1 & +5 & +10 & & \\
 & +2 & & & \\
\hline
1 & +7 & & &
\end{array}$$

The first transformed equation is

$$f_1(x_1) = x_1^3 + 7x_1^2 + 10x_1 - 1 = 0,$$

which has a root between 0 and 1. The root x_1 is small. Hence the first two terms are small compared to the last two. We obtain an approximate value, $x_1 = 0.1$, from $10x_1 - 1 = 0$. This value of x_1 makes the first two terms positive and $f_1(0.1) > 0$;

hence the constant term in the second transformed equation would be positive. This shows us that this value of x_1 is too large since $f_1(0) = -1$. The constant term in each transformed equation must retain the same sign as the constant term in the original equation. For $x_1 = 0.09$, $f_1(x_1)$ is found to be negative. Diminish the roots of $f_1(x_1) = 0$ by 0.09. We obtain

$$f_2(x_2) = x_2^3 + 7.27x_2^2 + 11.2843x_2 - 0.042571 = 0.$$

From $11.2843x_2 - 0.042571 = 0$, we obtain $x_2 = 0.003$. Hence we diminish the roots of $f_2(x_2) = 0$ by 0.003.

The work for these two steps is as follows:

$$\begin{array}{rrrl}
1 + 7 & +\ 10 & -\ 1 & \underline{|0.09} \\
0.09 & +\ \ 0.6381 & +\ 0.957429 & \\
\hline
1 + 7.09 & +\ 10.6381 & -\ 0.042571 & \\
0.09 & +\ \ 0.6462 & & \dfrac{0.04}{11.3} = 0.003 \\
\hline
1 + 7.18 & +\ 11.2843 & & \\
0.09 & & & \\
\hline
1 + 7.27 & & & \\
+\ 0.003 & +\ \ 0.021819 & +\ 0.033918357 & \underline{|0.003} \\
\hline
1 + 7.273 & +\ 11.306119 & -\ 0.008652643 & \\
+\ 0.003 & +\ \ 0.021828 & & \\
\hline
1 + 7.276 & +\ 11.327947 & & \\
+\ 0.003 & & & \\
\hline
1 + 7.279 & & &
\end{array}$$

Our third transformed equation is

$$x_3^3 + 7.279x_3^2 + 11.327947x_3 - 0.008652643 = 0.$$

From the last two terms we find

$$0.0007 < x_3 < 0.0008,$$

whence

$$0.000003570 < x_3^3 + 7.279x_3^2 < 0.000004664.$$

We may ignore the first two terms, provided the constant term 0.008652643 be reduced by an amount between these two limits. We have

$$0.008652643 - 0.000003570 = 0.008649073.$$

$$0.008652643 - 0.000004664 = 0.008647979.$$

From $11.327947x_3 - 0.008647979 = 0$, we obtain

$$x_3 = 0.0007634+.$$

From $11.327947x_3 - 0.008649073 = 0$, we obtain

$$x_3 = 0.0007635+.$$

Hence correct to 6 decimal places we have $x_3 = 0.0007634+$.

The work can be shortened as shown below:

$$
\begin{array}{r|r|l}
\underline{11.\dot{3}\dot{2}7947} & 0.008647 & \underline{0.0007634} \\
 & \underline{7929} & \\
 & 718 & \\
 & \underline{679} & \\
 & 39 & \\
 & \underline{34} & \\
 & 5 &
\end{array}
$$

Since the quotient is 0.0007, we use only two decimals in the divisor, except by inspection, to see how much should be carried in making the first multiplication. Place a dot $\cdot$ above the figure 2 in the divisor and use 11.32 as a divisor. Before multiplying by 6, the second significant figure in the quotient, place a dot over the figure 3 and use 11.3.

For the root of the original equation $f(x) = 0$, we have then

$$x = 2.0937634+.$$

The first 6 decimal places are correct. There is doubt as to whether the last figure should be a 4 or a 5.

If more decimals are required, it is not necessary to form a new transformed equation. We need only to revise the constant term in $f_3(x_3) = 0$, making use of our present better value of x_3.

This contracted method may be used after 3 or 4 decimals have been found. By it we are furnished, in addition to the decimals already obtained, with as many more decimals less one as there are decimals in the trial divisor, with doubt only as to the last decimal.

10.2 Graphical discussion of Horner's method. If $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$, ($a_0 > 0$, n an integer > 0, all coefficients real), then $y = a_{n-1}x + a_n$ is the tangent line to $y = f(x)$ at $(0, a_n)$.

In Horner's method, after each transformation, an approximate value of the root sought is found by neglecting the terms of second degree and higher, and setting the first degree term equal to the negative of the constant term. This is equivalent to finding the equation of the tangent line to the curve at the point where the curve crosses the y-axis and then finding where this tangent line crosses the x-axis. In Horner's method each approximate value is greater than the previous approximation and is less than the positive root desired.

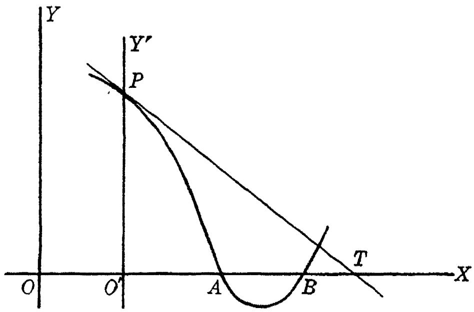


Fig. 22

The following questions arise: will the approximate value obtained, as outlined above in using the tangent line, be an approximate value to the desired root or to some other root of the equation? Will this be greater than the previous approximate values? Will this approximate value be less than the desired root?

Fig. 22 shows that if the desired root is in the interval OA, the point T of intersection of the tangent line PT may give an approximation to some root represented by the point B rather than to the desired root represented by the point A.

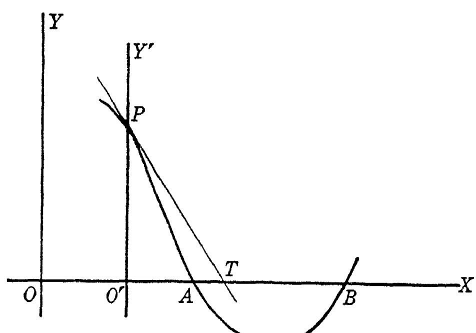


Fig. 23

Fig. 23 shows that the point T of intersection of the tangent line may give us a better approximation to the desired root than the point O' but that the value so obtained is too large and must be decreased.

Fig. 24 shows that changing the origin from O to O' increases the approximate value of the root and still keeps this approximate value less than the desired root. However, the next approxi-

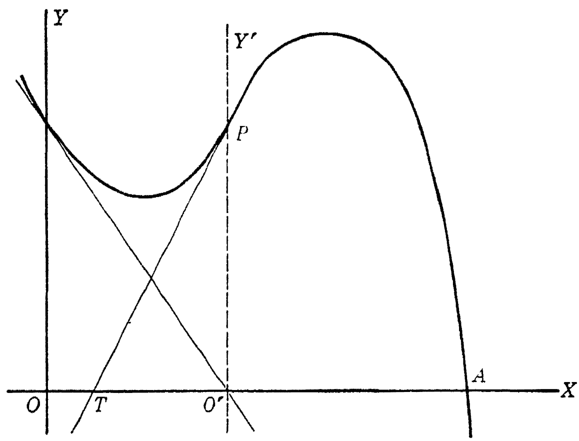


Fig. 24

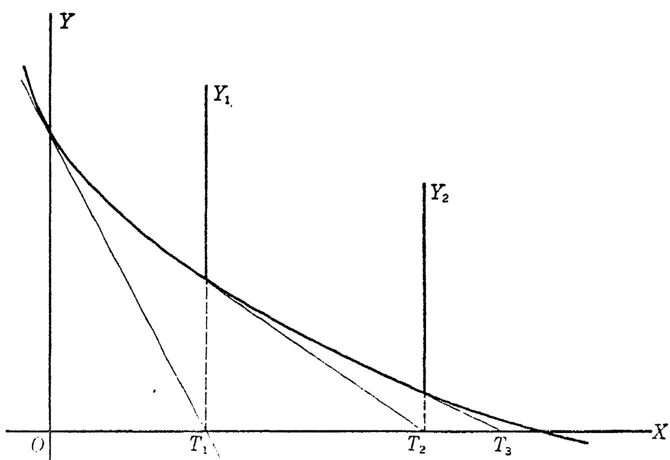


Fig. 25

mate value, represented by the point T, does not still further increase the approximate value and is not as good an approximation as that represented by the point O'.

Fig. 25 shows a case where each approximation, obtained by using the tangent line, is greater than the preceding, is less than the root desired, and is an approximation to the desired root and not to some other root.

Enough has been said to make it evident that great care must be exercised in using the point of intersection of the aforementioned tangent line with the x-axis to determine a better approximate value to the desired root.

Exercises

Verify the following:

1. The positive real root of $2x^3 - 17x^2 - 65x - 46 = 0$ is 11.5.
2. The positive real root of $4x^3 - 5x^2 - 43x - 34 = 0$ is 4.25.
3. The positive real root of $20x^3 + 19x^2 - 83x - 82 = 0$ is 2.05.
4. The positive real root of $x^3 + x^2 + x - 100 = 0$ is 4.2644 correct to four decimal places.
5. A root of $x^4 + 4x^3 - 4x^2 - 11x + 4 = 0$ is 1.6369 correct to four decimals.

Correct to 6 or 7 decimals the roots of the following equations are as indicated.

6. $x^4 - 8x^3 + 8x^2 - 8x + 3 = 0$	0.511304
7. $x^4 - x^3 + 2x^2 - 5x - 7 = 0$	2.041928
8. $x^4 - 12x + 7 = 0$	2.04727556; 0.59368583
9. $x^3 - 6x + 3 = 0$	0.523976; 2.145103
10. $x^3 - 7x + 7 = 0$	1.3568958; 1.6920215
11. $x^3 + 12x^2 + x - 100 = 0$	2.584455
12. $x^4 + 4x^3 - 8x - 4 = 0$	1.414214
13. $x^4 - 4x^3 + 6x^2 - 4x - 11 = 0$	2.861210
14. $x^3 - 6x^2 + 7 = 0$	1.208712
15. $x^4 - 12x^2 + 12x - 3 = 0$	2.858083; −3.907378
16. $x^3 - 2x^2 + 3x - 20 = 0$	3.106623
17. $x^5 + 12x^4 + 59x^3 + 150x^2 + 201x - 207 = 0$	0.6386058

18. Calculate to three decimal places each of the roots between 4 and 5 of $x^4 + 8x^3 - 70x^2 - 144x + 936 = 0$ 4.243; 4.246

19. A house may be bought for \$4400 cash, or for \$1000 cash and \$1000 at the end of each year for four years. Find the annual rate of interest implied in this offer.

After having paid \$1000 cash there remains \$3400 to be paid. The amount of \$3400 at the end of the four years should be equal to the sum of the amounts of the separate payments at the end of the four-year period. Hence, if x is the rate of interest,

$$3400(1 + x)^4$$
$$= 1000(1 + x)^3 + 1000(1 + x)^2 + 1000(1 + x) + 1000.$$

This reduces to $17x^4 + 63x^3 + 82x^2 + 38x - 3 = 0$

Ans.: $x = 0.068+$

20. A house may be bought for \$3800 cash, or for \$1000 cash and \$1000 at the end of each year for three years. Find the annual rate of interest implied in the offer. Ans.: 3.5%

21. The diameter of a water pipe whose length is 200 feet, and which is to discharge 100 cubic feet per second under a head of 10 feet, is given by the real root of the equation

$$x^5 - 38x - 101 = 0.$$

Find the diameter correct to 3 significant figures. Ans.: $x = 2.92$
[Merriman and Woodward, Higher Mathematics, p. 13.]

22. Let a sphere with specific gravity $g < 1$, be immersed in water. Let the radius of the sphere be r and the depth to which it will sink be h. Then the volume of the submerged portion is $\pi h^2\left(r - \frac{h}{3}\right)$. The volume of the sphere is $\frac{4}{3}\pi r^3$. The volume of the water displaced is $\frac{4}{3}\pi r^3 g$. Then

$$\pi h^2\left(r - \frac{h}{3}\right) = \frac{4}{3}\pi r^3 g,$$

or

$$h^3 - 3rh^2 + 4gr^3 = 0.$$

Find h for

(a) $r = 1, g = \frac{1}{4}$ $h = 0.65$
(b) $r = 1, g = \frac{3}{4}$ $h = 1.34$
(c) $r = 2, g = \frac{7}{8}$ $h = 3.11$

10.3 Newton's method. Newton's method can be applied to any equation $f(x) = 0$, algebraic or transcendental. Horner's method is applicable only to those equations in which $f(x)$ is a polynomial in x.

First determine two numbers a and b $(a < b)$ such that there is

one and only one root of $f(x) = 0$ between a and b. We now have an approximate value of the root.

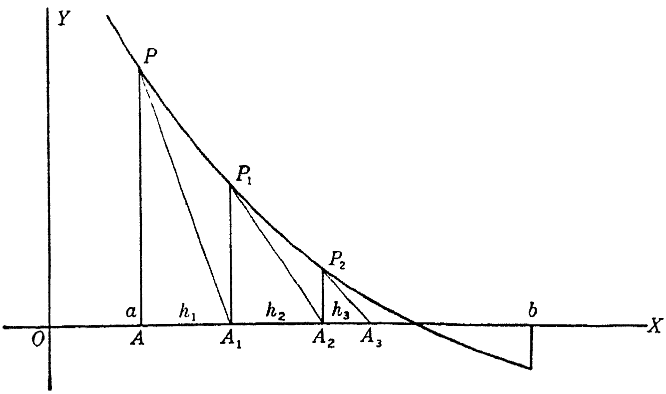


Fig. 26

If the curve is as represented in Fig. 26, we can find a correction h_1 as follows: let PA_1 be tangent to the curve at P. Then

$$\tan XA_1P = \left.\frac{dy}{dx}\right]_{x=a} = f'(a)$$
$$= -\frac{AP}{AA_1} = -\frac{f(a)}{h_1}.$$

Whence

$$h_1 = -\frac{f(a)}{f'(a)}.$$

Then $a_1 = a - \frac{f(a)}{f'(a)}$ is a second approximation to the root.

Obviously we can erect an ordinate to the curve from A_1, where $x = a + h_1 = a_1$. Let this meet the curve in P_1. Draw the tangent P_1A_2 at P_1. $A_1A_2 = h_2$ is a second correction. We have

$$\tan XA_2P_1 = \left.\frac{dy}{dx}\right]_{x=a_1} = f'(a_1) = -A_1P_1 = -\frac{f(a_1)}{h_2};$$

whence

$$h_2 = -\frac{f(a_1)}{f'(a_1)}.$$

Then $a_2 = a_1 - \dfrac{f(a_1)}{f'(a_1)}$ is a third approximation to the root.

One may begin as shown in Fig. 27 at $x = b$, at the upper end of the interval, and decrease the value $x = b$ successively, to

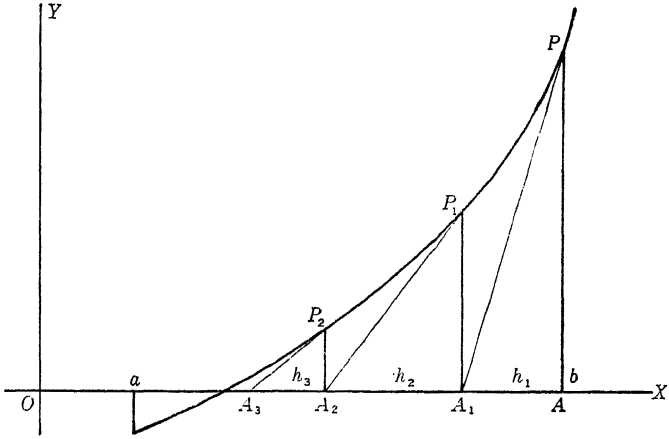


Fig. 27

obtain better and better approximations. The work would be as follows:

$$\tan XA_1P = \left.\frac{dy}{dx}\right]_{x=b} = f'(b)$$
$$= \frac{AP}{A_1A} = \frac{f(b)}{h_1}.$$

Whence

$$h_1 = \frac{f(b)}{f'(b)},$$

and

$$OA_1 = b_1 = b - h_1 = b - \frac{f(b)}{f'(b)} \text{ is a second approximation.}$$

$$OA_2 = b_2 = b_1 - h_2 = b_1 - \frac{f(b_1)}{f'(b_1)} \text{ is a third approximation.}$$

$$OA_3 = b_3 = b_2 \div \frac{f(b_2)}{f'(b_2)} \text{ is a fourth approximation, etc.}$$

Let us apply this method to the equation

$$f(x) = x^3 - 2x - 5 = 0.$$

$f(2) = -1; f(3) = +16$. Hence there is a root between 2 and 3. For the second approximation, we have

$$h_1 = -\frac{f(2)}{f'(2)} = +0.1,$$

whence

$$a_1 = 2 + 0.1 = 2.1$$

$$f(2.1) = +0.061.$$

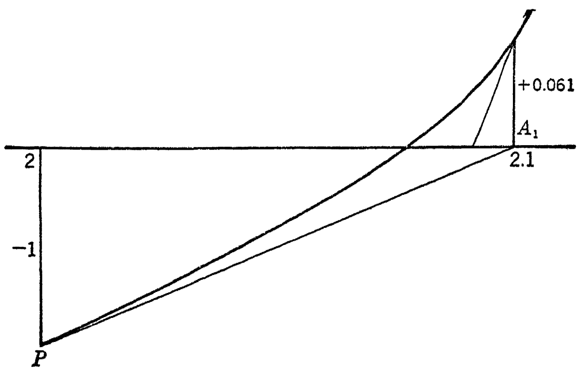


Fig. 28

Since $f(2.1)$ is positive, while $f(2)$ is negative, we know that the root is less than 2.1. For the third approximation, we have

$$h_2 = -\frac{f(2.1)}{f'(2.1)} = -\frac{0.061}{11.23} = -0.0054,$$

whence

$$a_2 = a_1 - \frac{f(a_1)}{f'(a_1)} = a_1 - 0.0054 = 2.0946.$$

For the next approximation, we have

$$a_3 = a_2 - \frac{f(a_2)}{f'(a_2)} = 2.0946 - 0.00004852 = 2.09455148$$

in which 7 decimal places are correct. We cannot be certain of the accuracy of the 8th decimal.

10.4 Graphical discussion of conditions for the success of Newton's method.

(A) $f'(x) = 0$ *between* a *and* b. In general one can begin the series of approximations at either end of the interval. However, if $f'(x) = 0$ for some point Q on the curve, $y = f(x)$, between a and b, one must not begin the approximation at a point P such that Q is between P and R; for if one begins at P, the second approximation OA_1 will bring one outside the interval a to b within which the root is situated. A continuation of the process

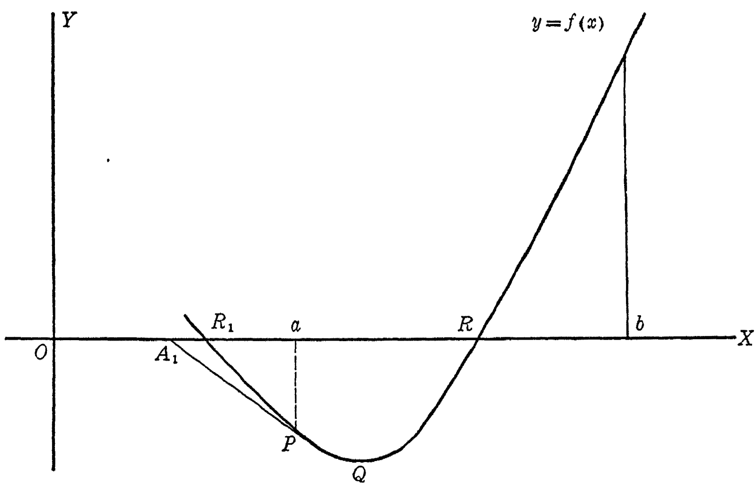


Fig. 29

from A_1 will often bring one to some other root R_1, but not to the root R which is sought.

(B) $f''(x) \neq 0$ between a and b.

Let $y = f(x)$ be represented by PIQ (Fig. 30). The slope of the curve is positive between P and Q. The slope increases from P to I and decreases from I to Q. Hence at I there is a maximum slope. This means that the ordinate CM on the first derivative curve $y = f'(x)$ is a maximum. I is said to be a *point of inflection* on the curve $y = f(x)$. Whenever a curve has a maximum (minimum) for a certain value of the abscissa, its first derivative vanishes, provided the first derivative is a continuous function. Then $f''(x) = 0$ for the abscissa of the point of inflection I. The

curve $y = f''(x)$ goes through the point C. To the left of C, $y = f''(x)$ is positive since to the left of M the slope of $y = f'(x)$ is positive. It might very well happen that either the tangent at P, or the tangent at Q meets the x-axis outside of the interval a to b. This might happen at only one end. It might not happen at all.

10.5 Fourier's theorem. *If $f(x) = 0$ be a rational integral algebraic equation which has one and only one real root between α and β, $\alpha < \beta$; and if $f'(x) = 0$ has no real root between α and β, and also $f''(x) = 0$ has no real root between α and β; then Newton's method of approximation will certainly be successful if it be begun*

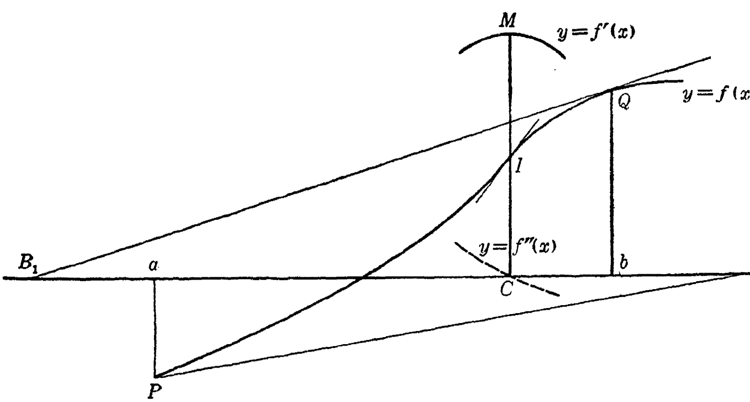


Fig. 30

and continued from that bound for which $f(x)$ and $f''(x)$ have the same sign.

It follows from our hypothesis that $f(x)$ changes sign only once in the interval and that $f'(x)$ and $f''(x)$ do not change sign in the interval.

Lemma: If $f(x)$ and $f'(x)$ are continuous in the interval a to b, then there exists a value λ of x between a and b for which

$$\frac{f(b) - f(a)}{b - a} = f'(\lambda).$$

Geometrically this lemma is obvious. What it says is that if the curve $y = f(x)$ is continuous and has a derivative which is continuous between a and b; then there is some point A on the curve, between P and Q, such that the slope of the secant line

PQ is the same as the slope of the tangent line at A. The equality may be written

$$f(b) - f(a) = (b - a)f'(\lambda).$$

Suppose $b > a$. Then it is easy to see from the last equation that if $f'(x)$ retains the same sign throughout the interval, then

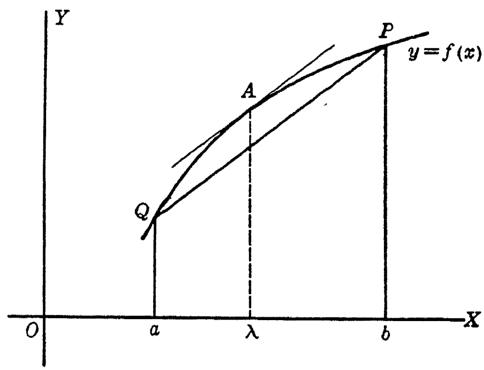


Fig. 31

as x increases throughout that interval $f(x)$ *increases numerically* when it has the same sign as $f'(x)$, and *decreases numerically* when it has the contrary sign.

Proof of Fourier's theorem:

(1) Suppose $f(x)$ and $f''(x)$ have the same sign when $x = \alpha$. Take α for the first approximation; then Newton's second approximation is $OA_1 = \alpha - \dfrac{f(\alpha)}{f'(\alpha)}$. Let $OR = \alpha + c$ represent the true value of the root; then $f(\alpha + c) = 0$. Now by the lemma, we have

$$f(\alpha + c) - f(\alpha) = cf'(\lambda), \text{ where } \lambda \text{ is between } \alpha \text{ and } \alpha + c;$$

thus $c = -\dfrac{f(\alpha)}{f'(\lambda)}$, and the true value of the root is

$$OR = \alpha + c = \alpha - \frac{f(\alpha)}{f'(\lambda)}.$$

We have to show that $\alpha - \dfrac{f(\alpha)}{f'(\alpha)}$ is nearer to the root than α. Since c is necessarily positive, $f(\alpha)$ and $f'(\lambda)$ are of different sign.

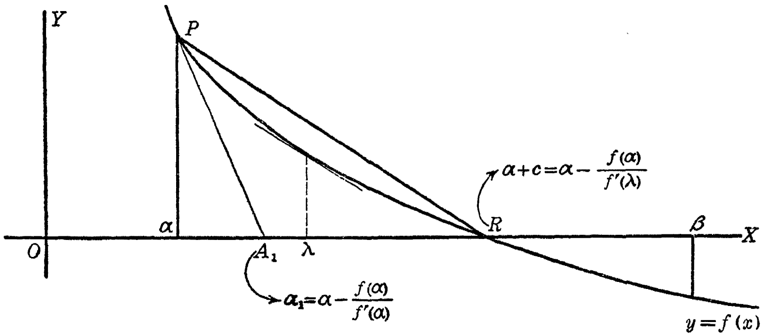


Fig. 32

By our hypotheses $f(\alpha)$ and $f''(\alpha)$ have the same sign; and therefore $f'(\lambda)$ and $f''(\alpha)$ are of contrary sign.

Hence, by our lemma, $f'(x)$ decreases numerically as x increases from α to β, so that $f'(\lambda)$ is numerically less than $f'(\alpha)$; therefore $-f(\alpha)/f'(\alpha)$ is a positive quantity which is numerically less than the positive quantity $-f(\alpha)/f'(\lambda)$. This shows that Newton's second approximation is nearer the true value of the root than α.

Let $\alpha_1 = \alpha - \dfrac{f(\alpha)}{f'(\alpha)}$; then $f(\alpha_1)$ and $f''(\alpha_1)$ have the same sign, and the approximation can be continued from α_1.

(2) Suppose that $f(x)$ and $f''(x)$ have the same sign when $x = \beta$. Let β be the first approximation, then Newton's second approximation is $\beta - \dfrac{f(\beta)}{f'(\beta)}$. Let $\beta + c$ represent the true value of the root. Then $f(\beta + c) = 0$. From the lemma we have

$$f(\beta + c) - f(\beta) = cf'(\lambda)$$

where λ is between β and $\beta + c$. Then $c = -\dfrac{f(\beta)}{f'(\lambda)}$. We have to show that $\beta - f(\beta)/f'(\beta)$ is nearer the true value of the root than β. Since c is necessarily negative, $f(\beta)$ and $f'(\lambda)$ are of the same sign, and since by hypothesis $f(\beta)$ and $f''(\beta)$ are of the same sign, then $f'(\lambda)$ and $f''(\beta)$ are of the same sign. Hence, by the lemma, $f'(x)$ increases numerically as x increases from α to β, and so $f'(\lambda)$ is less than $f'(\beta)$. Therefore

$$0 < \frac{f(\beta)}{f'(\beta)} < \frac{f(\beta)}{f'(\lambda)}.$$

This shows that Newton's second approximation is nearer the true value of the root than β.

Let $\beta_1 = \beta - \dfrac{f(\beta)}{f'(\beta)}$, then $f(\beta_1)$ and $f''(\beta_1)$ have the same sign, and the approximation can be continued from β_1.

This shows that Fourier's conditions are *sufficient*. We will now show that they are *necessary*.

If we start with a first approximation $x = \alpha$, Newton's second approximation corrects this by adding $-\dfrac{f(\alpha)}{f'(\alpha)}$, while the true value of the root is obtained by adding $-\dfrac{f(\alpha)}{f'(\lambda)}$. Hence the permanence of sign of $f'(x)$ is necessary in order that we may be sure that $f'(\alpha)$ and $f'(\lambda)$ have the same sign. If these equations do not have the same sign, the Newtonian correction has the wrong sign, and Newton's second approximation is further from the true value of the root tnan the first approximation. The permanence of sign of $f''(x)$ is necessary in order to ensure that $f'(\lambda)$ is numerically less than $f'(\alpha)$. If this is not the case, the Newtonian correction is greater than the true correction, and so, supposing the correction to be in the right direction, the true value of the root is between Newton's first and second approximations. In this case Newton's second approximation may be nearer the true value of the root than the first approximation, but is not necessarily so.

Exercises

Compute one root of each of the following equations:

1. $x^3 + 3x - 5 = 0$
2. $x^4 - 4x^3 - 4x + 12 = 0$
3. $x^3 - 3x^2 + 3 = 0$
4. $x^3 - 6x^2 + 19 = 0$
5. $x^3 - x^2 - 2x + 1 = 0$
6. $x^5 - x^4 - 4x^3 + 3x^2 + 3x - 1 = 0$
7. $x^3 + x^2 + x - 100 = 0$
8. $20x^3 + 24x^2 - 3 = 0$
9. $\cos x - x = 0$
10. $\sin 2x - x = 0$
11. $2 \sin x - x = 0$
12. $e^x - \tan x = 0$
13. $e^x - \sin x = 0$
14. $2 \sin x = (x - 1)^2$
15. $\tan x + x - 1 = 0$

Compute each real root of the following equations correct to two decimals:

16. $x^3 - 6x^2 + 3x + 5 = 0$
17. $x^4 - 4x^3 - 4x + 12 = 0$
18. $x^3 - 9x - 5 = 0$

10.6 Graeffe's method. By this method one finds all of the roots, real and complex, without any preliminary determination of their approximate values. The method consists in forming a new equation whose roots are some high power of the roots of the given equation. We discuss first the case where all roots are *real.*

As the power increases the distance between the roots increases. For high powers, roots whose numerical value is greater than 1 become *widely separated.* We shall see that an equation whose real roots are widely separated can be readily solved.

Let the equation to be solved be

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n$$
$$= 0 \text{ (coefficients real).} \qquad (1)$$

Write (1) so that all terms of even degree are on one side of the equation and all terms of odd degree are on the other side. Square both sides. We have

$$(x^n + a_2x^{n-2} + a_4x^{n-4} + \cdots)^2 = (a_1x^{n-1} + a_3x^{n-3} + \cdots)^2.$$

Put $x^2 = -y$. We have

$$y^n + b_1y^{n-1} + b_2y^{n-2} + \cdots + b_{n-1}y + b_n = 0 \qquad (2)$$

where

$$b_1 = a_1^2 - 2a_2$$

$$b_2 = a_2^2 - 2a_1a_3 + 2a_4$$

$$b_3 = a_3^2 - 2a_2a_4 + 2a_1a_5 - 2a_6$$

$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$

$$b_r = a_r^2 - 2a_{r-1}a_{r+1} + \cdots + (-1)^s 2a_{r-s}a_{r+s} + \cdots + (-1)^r 2a_{2r}$$

$$b_n = a_n^2.$$

A rule for the formation of the coefficients in (2) may be stated as follows: *the coefficient of any power of y is formed by squaring the coefficient of the corresponding power of x in the original equation and adding twice the product of every pair of coefficients which are equally distant on either side, these products being taken with signs alternately positive and negative, missing powers of x being supplied with zero coefficients.*

Suppose the roots of (1) to be $-a$, $-b$, $-c$, $\cdots$ (all real) and $|a| > |b| > |c| > \cdots$. Then a, b, c, $\cdots$ are called "Encke roots" of (1). Now since $y = -x^2$, the Encke roots of (2) are a^2, b^2, c^2, $\cdots$.

Repeat the process, thus obtaining an equation whose Encke roots are the squares of the roots of (2), that is, the 4th powers of the roots of (1). At the next step one obtains an equation whose Encke roots are the 8th powers of the roots of (1), and so on. Perform this operation k times. Let $2^k = m$. The equation whose Encke roots are a^m, b^m, c^m, $\cdots$ is

$$(x + a^m)(x + b^m)(x + c^m) \cdots = 0,$$

or

$$x^n + [a^m]x^{n-1} + [a^m b^m]x^{n-2} + [a^m b^m c^m]x^{n-3} + \cdots = 0,$$

where

$$[a^m] = a^m + b^m + c^m + \cdots$$

$$[a^m b^m] = a^m b^m + a^m c^m + \cdots + b^m c^m + \cdots, \text{ etc.}$$

Continue this process until the doubled products bring no change in the digits we wish to retain.

Since $|a| > |b| > |c| > \cdots$, if m is sufficiently large the ratio of a^m to $[a^m]$ is approximately one. Likewise the ratio of $a^m b^m$ to $[a^m b^m]$ is approximately one.

We can now determine the numerical value of the roots, a, b, c, $\cdots$ in succession from the equations

$$[a^m] = a^m$$

$$[a^m b^m] = a^m b^m$$

$$[a^m b^m c^m] = a^m b^m c^m, \text{ etc.}$$

Descartes' rule of signs, a graph, substitution in the original equation will help in the determination as to whether a root is positive or negative.

Example. Find the roots of $x^3 + 6x^2 + 11x + 6 = 0$.

Use the notation $7.654 \times 10^8 = 7.654^8$.

The work is arranged as shown on page 150.

1	1	6	11	6
	1	36 −22	121 −72	36
2	1	14	49	36
	1	196 −98	2401 −1008	1296
4	1	98	1393	1296
	1	9.604^3 −2.786	1.940^6 −0.254	1.680^6
8	1	6.818^3	1.686^6	1.680^6
	1	4.649^7 −0.337	2.843^{12} −0.023	2.822^{12}
16	1	4.312^7	2.820^{12}	2.822^{12}
	1	1.859^{15} −0.006	7.952^{24} −0.000	7.964^{24}
32	1	1.853^{15}	7.952^{24}	7.964^{24}

$$a^{32} = 1.853^{15};\ \log a^{32} = 15.2679;\ \log |a| = 0.4771;\ |a| = 3.000$$

$$a^{32}b^{32} = 7.952^{24};\ \log b^{32} = 24.9005 - 15.2679 = 9.6326$$

$$\log |b| = 0.3010;\ |b| = 2.000$$

$$(abc)^{32} = 7.964^{24};\ \log c^{32} = 24.9011 - 24.9005 = 0.0006$$

$$\log |c| = 0.0000;\ |c| = 1.000.$$

By Descartes' rule of signs there are no positive roots. The roots are -1, -2, -3.

10.7 Graeffe's method. Complex roots.

I: *One pair of complex roots.*

If some roots are complex, the method of the preceding article can be used to determine their values. Let us illustrate the method by the simplest case, that of a cubic with one real root

and two complex roots. Let the Encke roots of the given equation be

$$a, \; r(\cos\theta + i\sin\theta), \qquad r(\cos\theta - i\sin\theta), \qquad r > 0.$$

The equation whose Encke roots are the m th powers of these is

$$(x + a^m)[x + r^m(\cos m\theta + i\sin m\theta)][x + r^m(\cos m\theta - i\sin m\theta)] = 0,$$

or

$$x^3 + (a^m + 2r^m\cos m\theta)x^2 + (r^{2m} + 2a^m r^m \cos m\theta)x + a^m r^{2m} = 0. \quad (3)$$

If $|a| > r$, and m is large enough, a^m is large compared to $2r^m \cos m\theta$ and a can be computed by taking the m th root of the coefficient of x^2. Then r can be computed from the constant term $a^m r^{2m}$.

If $|a| < r$, then r^{2m} is large compared to $2a^m r^m \cos m\theta$ and r can be computed by taking the $2m$th root of the coefficient of x. Then a can be computed from the constant term.

Suppose the complex roots are $u \pm iv$. We can compute u since the sum of the roots satisfies the relation

$$-a_1 = 2u - a;$$

then v can be computed from the relation $r^2 = u^2 + v^2$.

If $|a| > r$, and m is large, equation (3) becomes essentially

$$x^3 + a^m x^2 + 2a^m r^m \cos m\theta x + a^m r^{2m} = 0.$$

For real roots, the coefficients, for m large, are to a close degree of approximation the squares of the coefficients in the preceding step. That is, all coefficients are positive. If there is a pair of complex roots, for the cubic, the coefficient of x, owing to the presence of the factor $\cos m\theta$, will fluctuate in sign.

If $|a| < r$, and m is large, equation (3) becomes essentially

$$x^3 + 2r^m \cos m\theta x^2 + r^{2m}x + a^m r^{2m} = 0.$$

If there is a pair of complex roots, the coefficient of x^2 will fluctuate in sign.

II. *Two pairs of complex roots.* Since the presence of real roots neither simplifies nor complicates the procedure for finding the complex roots, we may confine our discussion to equations all of whose roots are imaginary.

Let the Encke roots of the given quartic be

$$r(\cos\phi \pm i\sin\phi), \quad s(\cos\theta \pm i\sin\theta), \quad r > 0, \quad s > 0.$$

The equation whose Encke roots are the m th powers of these is

$$\begin{aligned} x^4 &+ 2(r^m \cos m\phi + s^m \cos m\theta)x^3 \\ &+ (r^{2m} + 4r^m s^m \cos m\phi \cos m\theta + s^{2m})x^2 \\ &+ 2r^m s^m(r^m \cos m\theta + s^m \cos m\phi)x + r^{2m}s^{2m} = 0. \end{aligned}$$

The approximate equations are:
If $r > s$,

$$x^4 + 2r^m \cos m\phi x^3 + r^{2m}x^2 + 2r^{2m}s^m \cos m\theta\cdot x + r^{2m}s^{2m} = 0.$$

r can be determined from the coefficient of x^2 and then s from the constant term.
If $r < s$,

$$x^4 + 2s^m \cos m\theta x^3 + s^{2m}x^2 + 2r^m s^{2m} \cos m\phi\cdot x + r^{2m}s^{2m} = 0.$$

s can be determined from the coefficient of x^2 and then r from the constant term.
Now let the complex roots of the original equation be represented by $u_1, \pm iv_1$ and $u_2 \pm iv_2$. The equation, with real coefficients,

$$f(x) \equiv x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0,$$

to be solved may be written

$$\begin{aligned} f(x) \equiv x^4 - 2(u_1 + u_2)x^3 &+ (r_1^2 + r_2^2 + 4u_1u_2)x^2 \\ - 2(u_1r_2^2 &+ u_2r_1^2)x + r_1^2r_2^2 = 0, \\ r_1^2 = u_1^2 + v_1^2, &\qquad r_2^2 = u_2^2 + v_2^2. \end{aligned} \tag{4}$$

We can now determine u_1, u_2 from the equations

$$\begin{aligned} 2u_1 + 2u_2 &= -a_1 \\ 2r_2^2u_1 + 2r_1^2u_2 &= -a_3. \end{aligned}$$

Having found u_1 and u_2, we determine v_1 and v_2 from (4).

III. *Three pairs of complex roots.* The equation which has the three pairs of roots

$$u_j \pm iv_j, \qquad j = 1, 2, 3$$

is

$$\begin{aligned} x^6 &- 2[u_1 + u_2 + u_3]x^5 + [r_1^2 + r_2^2 + r_3^2 + 4u_1u_2 + 4u_1u_3 + 4u_2u_3]x^4 \\ &- [2u_1(r_2^2 + r_3^2) + 2u_2(r_1^2 + r_3^2) + 2u_3(r_1^2 + r_2^2) + 8u_1u_2u_3]x^3 \\ &+ [r_1^2r_2^2 + r_1^2r_3^2 + r_2^2r_3^2 + 4(u_1u_2r_3^2 + u_1u_3r_2^2 + u_2u_3r_1^2)]x^2 \\ &- 2[r_2^2r_3^2u_1 + r_1^2r_3^2u_2 + r_1^2r_2^2u_3]x + r_1^2r_2^2r_3^2 = 0. \end{aligned}$$

After we have found r_1, r_2, r_3 by Graeffe's method, if we are to proceed as we did in the case of two complex roots, we must determine u_1, u_2, u_3 from the equations

$$\begin{aligned} 2u_1 + 2u_2 + 2u_3 &= -a_1 \\ 2(r_2^2 + r_3^2)u_1 + 2(r_1^2 + r_3^2)u_2 + 2(r_1^2 + r_2^2)u_3 + 8u_1u_2u_3 &= -a_3 \\ 2r_2^2r_3^2u_1 + 2r_1^2r_3^2u_2 + 2r_1^2r_2^2u_3 &= -a_5. \end{aligned}$$

Elimination of u_2, u_3 leads to a cubic in u_1 which can be solved.

For four or more pairs of complex roots the problem becomes more difficult and will not be taken up in this course.

Example. Find the roots of

$$x^5 - x^4 - x^3 + 19x^2 - 32x + 30 = 0.$$

The detail of the work is not shown. The transformed equations are arranged in the following table.

	x^5	$-x^4$	$-x^3$	$+19x^2$	$-32x$	$+30$
1	1	-1	-1	$+19$	-32	30
2	1	3	-25	237	-116	900
4	1	59	-1.0290^3	5.5769^4	-4.1314^5	8.1000^5
8	1	5.5390^3	-6.3482^6	2.3556^9	8.0330^{11}	6.5610^{11}
16	1	4.3377^7	1.4365^{13}	6.5761^{18}	3.3620^{21}	4.3047^{23}
32	1	1.8519^{15}	-3.6414^{26}	4.3149^{37}	0.5641^{43}	1.8630^{47}
64	1	3.4302^{30}	-2.7220^{52}	1.8618^{75}	1.5744^{85}	3.4708^{94}

$$a^{64} = 3.4302^{30} \text{ whence } |a| = 3, \quad a = -3$$

$$a^{64}r_1^{128} = 1.8618^{75} \text{ whence } r_1^2 = 5$$

$$a^{64}r_1^{128}r_2^{128} = 3.4708^{94} \text{ whence } r_2^2 = 2.$$

The roots are -3, $u_1 \pm iv_1$, $u_2 \pm iv_2$. The sum of the roots $= 1$. There are two equations to determine u_1, u_2:

$$2u_1 + 2u_2 - 3 = 1 \quad \text{and} \quad -6(u_2r_1^2 + u_1r_2^2) = -32$$

whence

$$u_1 = u_2 = 1, \quad v_1 = \sqrt{r_1^2 - 1} = 2, \quad v_2 = \sqrt{r_2^2 - 1} = 1$$

The roots are -3, $1 \pm i$, $1 \pm 2i$.

10.8 Graeffe's method. Equal roots. There is no need to use Graeffe's method for this case. If the equation has equal roots, these can be detected and eliminated by finding the H.C.F. of $f(x)$ and $f'(x)$. However, if one has not made this test and elimination initially, we will now proceed to explain how the presence of equal roots will become evident as one progresses with the systematic computation by Graeffe's method.

Let the Encke roots of the given equation $f(x) = 0$ be

$$a, b, b, c, d, \cdots \quad \text{where} \quad |a| > |b| > |c| > |d| > \cdots.$$

The equation whose Encke roots are the m th powers of the roots of $f(x) = 0$ is

$$x^n + [a^m]x^{n-1} + [a^mb^m]x^{n-2} + [a^mb^mb^m]x^{n-3} + \cdots = 0$$

The coefficient $[a^mb^m] = 2a^mb^m + a^mc^m + \cdots$.

This coefficient, for m large, differs but little from $2a^mb^m$. Ordinarily when m is doubled, this coefficient, for m large, is approximately squared. In this case, however, there comes a place in the computation where at each step this coefficient is one half the square of the corresponding coefficient in the previous step, for

$$\tfrac{1}{2}(2a^mb^m)^2 = 2a^{2m}b^{2m}.$$

Exercises

By Graeffe's method find the roots of *Ans.*

1. $x^3 - 11x + 150 = 0$ — $-6, 3 \pm 4i$
2. $x^4 - 10x^3 + 39x^2 - 70x + 50 = 0$ — $2 \pm i, 3 \pm i$
3. $x^3 + 8x^2 + 21x + 18 = 0$ — $-3, -3, -2.$

Find the largest real root of the following equations correct to two decimal places.

	Ans.
4. $x^3 - 9x - 5 = 0$	3.25
5. $x^3 - 6x^2 + 3x + 5 = 0$	5.25
6. $x^4 - 4x^3 - 4x + 12 = 0$	4.06
7. $x^3 - 2x - 2 = 0$	1.77
8. $x^4 - 4x - 2 = 0$	1.73

Find the values of the real roots of the following equations correct to two decimal places.

9. $2x^3 - 4x^2 - 10x - 3 = 0$	3.53; −1.17; −0.36
10. $8x^3 - 12x^2 + 1 = 0$	1.44; 0.33; −0.27
11. $5x^3 - 3x^2 - 6x + 3 = 0$	1.18; 0.48; −1.06

References

Whittaker and Robinson: *Calculus of Observations*, London, 1924.
Hutchinson, C. A.: *Am. Math. Monthly*, 1935, pp. 149–161.
Hausmann, B. A.: *Am. Math. Monthly*, 1936, pp. 225–229.

Chapter XI

DETERMINANTS

11.1 Introduction. In Chapter II we discussed determinants of the second and third order and their application to the solution of simultaneous linear equations in two and three variables. In this chapter we will discuss the solution of simultaneous linear equations in n unknowns. The solutions can be expressed symbolically in terms of determinants of order n.

For determinants of order two and three there is a diagrammatic method for obtaining the number for which the determinant stands. No diagrammatic method exists for determining the number for which a determinant of order $n(n>3)$ stands. In this chapter there will be developed methods of handling determinants that will assist us in determining the number for which the determinant stands.

11.2 Permutations. In elementary algebra, it is proven that the number of permutations, or arrangements, of n things taken n at a time is $n!$ (read "n factorial"). For example, the $3! = 6$ permutations of the three letters a, b, c are as follows:

$$abc, acb, bac, bca, cab, cba$$

If the members of a group* are represented by letters of the Latin alphabet, then the group is said to be in its natural order if the letters occur in alphabetical sequence. Thus the group

$$a \quad g \quad k \quad n \quad p \quad s \quad x$$

is in its natural order. If the group is represented by numbers, then the group is in its natural order if these numbers are arranged in order of magnitude, the smallest first. Thus the group

$$3, 7, 13, 26, 51, 83, 128$$

* The word "group" is here used in its nontechnical sense for a collection or set. The word "group" has also in algebra a special meaning with which we are not concerned here.

is in its natural order. Greek letters remain merely symbols representing numbers or Latin letters in some arrangement.

If in a given arrangement of a group any two members are arranged in their natural order, whether or not they occur in adjacent positions, they are said to constitute a *permanence*, otherwise an *inversion*. Thus the pairs

$$ab;\ pq;\ ar;\ bx;\ 1,\ 7;\ 2,\ 8;\ 5,\ 11$$

are permanences, while the pairs

$$7,\ 5;\quad 3,\ 2;\quad 7,\ 1;\quad 101,\ 87;\quad b,\ a;\quad n,\ c;\quad r,\ k;\quad a_3,\ a_1;\quad b_1,\ a_1$$

are inversions.

The number of permanences, or inversions, of the members of a group can be found without duplication by comparing each member with every member that comes after it. For example, in $a\ e\ g\ p\ c\ b$ the permanences are

$$ae,\ ag,\ ap,\ ac,\ ab,\ eg,\ ep,\ gp$$

and the inversions are

$$ec,\ eb,\ gc,\ gb,\ pc,\ pb,\ cb$$

In the permutation 5 4 1 2 3 there are 3 permanences and 7 inversions.

A permutation is said to be positive, or even, if it has an even number of inversions. For example 3 4 2 5 1 and $a\ e\ g\ b\ c$ have 6 and 4 inversions respectively.

A permutation is said to be negative, or odd, if it has an odd number of inversions. For example 5 4 1 2 3 has 7 inversions and $a\ e\ g\ c\ b$ has 5 inversions.

11.3 Theorem I: *If in any permutation, two members of the group are interchanged, the permutation is changed either from odd to even, or from even to odd.*

1. Suppose that the members to be interchanged are adjacent. Represent the permutation by

$$A\ \alpha\ \beta\ Z$$

where A represents the sequence of all of the members that come before α and Z represents that of all of the members that come after β, while α and β are the adjacent members. Now interchanging α and β in no way affects the groups A and Z or the

relation of α and β to the members of the groups A and Z. The only change is replacing $\alpha\beta$ by $\beta\alpha$. This makes but one change in the number of inversions. For if $\alpha\beta$ is a permanence, it is replaced by $\beta\alpha$, which is an inversion, and the number of inversions is increased by one. While if $\alpha\beta$ is an inversion, it is replaced by $\beta\alpha$, which is a permanence, and the number of inversions is diminished by one. Thus the permutation if even becomes odd, and if odd becomes even.

2. Suppose the members to be interchanged are not adjacent. Represent the permutation by

$$A \ \alpha \ N \ \beta \ Z$$

where N represents the members of the permutation between α and β. Let n be the number of the members of the group in N. Interchange β in turn with each of the n members of the group N. Then with n changes in the number of inversions we arrive at the order

$$A \ \alpha \ \beta \ N \ Z$$

Now interchange α in turn with each of the $n + 1$ members of the group (βN). Then with $n + 1$ additional changes in the number of inversions we arrive at the order

$$A \ \beta \ N \ \alpha \ Z$$

In all there have been

$$n + (n + 1) = 2n + 1$$

changes in the number of inversions. Since n is necessarily an integer, $2n + 1$ is an odd number. Hence the permutation has been changed from odd to even or from even to odd, and the theorem is proved.

Remark. This theorem may be also stated as follows: *If in any permutation two members of the group are interchanged, the number of inversions is changed by an odd number.*

11.4 Theorem II: *Of all the possible permutations of n things one half are odd and one half are even.* $(n > 1)$

Arrange all of the permutations in two sets. In the first set place all of the even permutations and in the second set place all of the odd permutations. Now interchange any two members. Every permutation in the first set now becomes odd and every

permutation in the second set becomes even. Taken as a whole we have the same permutations that we had originally, but arranged differently. That is, the number in the first set must be exactly equal to the number in the second set. Every member of the first set has moved over to the second set, but we have the same number of odd permutations in the second set that we had originally, although they may be differently arranged.

11.5 Theorem III: *If a permutation of the integers* $1, \cdots, n$, *can be obtained from the natural order, or can be restored to the natural order, by an even (odd) number of interchanges of a pair of numbers, it will have an even (odd) number of inversions.*

The natural order has no inversions. Each interchange of a pair of numbers introduces an odd number of inversions. Then an even number of interchanges will bring about an even number of inversions since an even number of odd numbers is even; and an odd number of interchanges will bring about an odd number of inversions since an odd number of odd numbers is odd.

11.6 Definition of a determinant of order n. A square array of elements consisting of n rows and n columns is called a square matrix of order n, and a certain rational function (to be described presently) of these elements is called a *determinant of order n.* Taking n letters

$$a, b, c, \cdots, l,$$

we may exhibit the elements of a determinant of order n and indicate the determinant as follows:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & \cdots & l_1 \\ a_2 & b_2 & c_2 & \cdots & l_2 \\ a_3 & b_3 & c_3 & \cdots & l_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & c_n & \cdots & l_n \end{vmatrix} \qquad (1)$$

Another notation commonly used is as follows:

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \qquad (2)$$

wherein the first subscript refers to the row in which the element is located and the second subscript refers to the column. When this notation is used, one should not, of course, read a_{11} as "*a subscript eleven*," but as "*a subscript one, one*," or more conveniently as "*a, one, one.*"

Write down all of the products that can be formed by taking as factors one, and only one, element from each row and each column. There will be $n!$ such products. Of these products one half involves (in a sense to be described) the even permutations, and one half involves the odd permutations.

In order to determine the sign of a particular product of the expansion of (1) arrange the factors in alphabetical order and then count the number of inversions of subscripts. The sign of the term is positive if the number of inversions is even and negative if the number of inversions is odd.

If we are considering the expansion of (2), the rule for determining the sign of a term becomes: arrange the factors so that the subscripts that indicate the row are in their natural order, then the sign of the term is positive (negative) if the number of inversions on the subscripts that indicate the column is even (odd).

The sequence of elements $a_1b_2c_3 \cdots l_n$ is called the principal diagonal of (1), and the term $a_1b_2c_3 \cdots l_n$ is called the principal diagonal term of the expansion. In this term the letters are in their natural order, the subscripts are also in their natural order, and the term is positive. All other terms in the expansion of (1) can be obtained from this principal diagonal by permuting the subscripts.

Example. Determine the sign of $b_4a_3d_1c_2e_5$. Rearrange so that the letters are in their natural order. We have $a_3b_4c_2d_1e_5$. Now count the inversions in the subscripts 3 4 2 1 5. There are five inversions; hence the term is negative.

Exercises

1. In a determinant of the fourth order, find the signs of the terms:

$$a_1b_3c_4d_2; \quad a_2b_1c_4d_3; \quad a_4b_1c_3d_2; \quad a_3b_1c_4d_2$$

2. In a determinant of the fifth order, find the signs of the terms:

$$a_1b_3c_5d_2e_4; \quad a_2b_4c_1d_3e_5; \quad a_1c_2b_4e_5d_3; \quad d_5e_3a_2c_1b_4$$

3. In determinant (4) of order 4, write out all of the terms with their proper signs that contain a_{14}.

4. If $a_1b_2c_3d_4e_5$ is one term of the determinant (1) of order 5, write out all of the terms with their proper signs that contain $a_2b_1c_4$.

5. If $a_1b_2c_3d_4e_5f_6$ is one term of the determinant (1) of order 6, write out all of the terms with their proper signs that contain $a_1b_2c_3$.

Properties of Determinants

11.7 Theorem IV: *The value of a determinant is not changed if the rows are changed into corresponding columns and the columns into corresponding rows; or, more briefly, if rows and columns are interchanged.*

Consider the determinant

$$D \equiv \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \tag{3}$$

If rows and columns are interchanged, we have

$$D_1 = \begin{vmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{n2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{n3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{vmatrix} \tag{4}$$

We observe that D_1 can be obtained from D by replacing a_{ij} by a_{ji}. A term of D is

$$(-1)^i a_{1c_1} a_{2c_2} a_{3c_3} \cdots a_{nc_n} \tag{5}$$

The first subscripts are in their natural order. The second subscripts

$$c_1, c_2, c_3, \cdots, c_n \tag{6}$$

are the integers $1, 2, \cdots, n$, but in an arbitrary order. The series (6) has, let us say, i inversions and hence the sign in (5) is $(-1)^i$.

By the substitution of a_{ji} for a_{ij}, (5) becomes

$$(-1)^i a_{c_1 1} a_{c_2 2} a_{c_3 3} \cdots a_{c_n n} \tag{7}$$

But (7) is a term of the expansion of D_1 and $(-1)^i$ is the proper sign. For the sign, by definition, is determined by the number of inversions on the subscripts which indicate the column when the

subscripts that indicate the row are in their natural order. Now in (7) it is the second subscript that indicates the row, and these subscripts are in their natural order, while the first subscripts indicate the column and these subscripts are the same set (6), in the same order that they appear in (5). There being by assumption i inversions in the set (5), the sign in (7) must be $(-1)^i$, the same as in (5). Hence every term that appears in D appears in D_1 and with the same sign.

11.8 Interchange of two columns. Theorem V: *If any two columns of a determinant are interchanged, the sign of the determinant is changed.*

Let Δ_1 be the determinant obtained from Δ by the interchange of the i-th and j-th columns. The terms of Δ_1 can be obtained from those of Δ by interchanging the subscripts i and j. By theorem I this changes the sign of every term. Hence $\Delta_1 = -\Delta$.

11.9 Interchange of two rows. Theorem VI: *If any two rows of a determinant are interchanged the sign of the determinant is changed.*

Let $\mathfrak{D}$ be obtained from Δ by changing columns into rows, then $\mathfrak{D} = \Delta$. Let Δ_1 be obtained from Δ by an interchange of two columns, say the i-th and j-th, then $\Delta = -\Delta_1$. Let $\mathfrak{D}_1$ be obtained from Δ_1 by interchanging columns and rows, then $\mathfrak{D}_1 = \Delta_1$. Whence $\mathfrak{D} = -\mathfrak{D}_1$ and $\mathfrak{D}_1$ can be obtained from $\mathfrak{D}$ by an interchange of the i-th and j-th rows.

11.10 Two rows (columns) alike. Theorem VII: *If two rows (columns) are alike the value of the determinant is zero.*

For by interchanging the two like rows (columns), the determinant is unchanged in value, but must change in sign. Hence

$$D = -D$$

whence

$$D = 0.$$

11.11 Theorem VIII: *Multiplying each element of a column (row) of a determinant by a given factor multiplies the determinant by that factor; that is,*

$$\begin{vmatrix} ma_1 & b_1 & \cdots & l_1 \\ ma_2 & b_2 & \cdots & l_2 \\ \cdot & \cdot & \cdot & \cdot \\ ma_n & b_n & \cdots & l_n \end{vmatrix} = m \begin{vmatrix} a_1 & b_1 & \cdots & l_1 \\ a_2 & b_2 & \cdots & l_2 \\ \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & \cdots & l_n \end{vmatrix}$$

This follows immediately from the definition of a determinant. Every term in the expansion contains one and only one element from the column (or row) in question. If each element of this column is multiplied by a given factor, each term of the expansion will be multiplied by the same factor; and hence, the determinant will be multiplied by the given factor.

Corollary I: *If each element of a column (row) of a determinant is zero, the determinant vanishes.*

Corollary II: *If the signs of all the elements in any column (row) are changed, the sign of the determinant is changed.* For this is equivalent to multiplying by the factor -1.

Exercises

1. Show that

$$\begin{vmatrix} 2 & 6 & 6 & 1 \\ 1 & 4 & 3 & 2 \\ 3 & 2 & 9 & 4 \\ 4 & 1 & 12 & 3 \end{vmatrix} = 0; \qquad \begin{vmatrix} 9 & 8 & 0 & 7 & 5 \\ 3 & 0 & 1 & 5 & 6 \\ 2 & 1 & 4 & 3 & 2 \\ 8 & 7 & 6 & 5 & 1 \\ 4 & 2 & 8 & 6 & 4 \end{vmatrix} = 0$$

2. Show that

$$\begin{vmatrix} 2 & 6 & 1 & 2 \\ 4 & 12 & 1 & 0 \\ 6 & 3 & 0 & 1 \\ 10 & 9 & 2 & 3 \end{vmatrix} = 6 \begin{vmatrix} 1 & 2 & 1 & 2 \\ 2 & 4 & 1 & 0 \\ 3 & 1 & 0 & 1 \\ 5 & 3 & 2 & 3 \end{vmatrix}$$

3. Show that

$$\begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

Hint: Multiply the rows by a, b, c respectively and divide the first column by abc.

4. Show that

$$\begin{vmatrix} bcd & a & a^2 & a^3 \\ cda & b & b^2 & b^3 \\ dab & c & c^2 & c^3 \\ abc & d & d^2 & d^3 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 & a^4 \\ 1 & b^2 & b^3 & b^4 \\ 1 & c^2 & c^3 & c^4 \\ 1 & d^2 & d^3 & d^4 \end{vmatrix}$$

5. Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-c)(c-a)(a-b)$$

If $a = b$, the determinant vanishes; hence by the factor theorem (§5.2), $a - b$ is a factor. In like manner $c - a$ and $b - c$ are factors. The principal diagonal is $+bc^2$. Arrange the signs of the factors so that $+bc^2$ is one term of the expansion.

6. Show that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix} = (c - b)(a - d)(c - a)(b - d)(a - b)(c - d)$$

7. Show that

$$\begin{vmatrix} 4 & 2 & 5 & 10 \\ 1 & 1 & 6 & 3 \\ 7 & 3 & 0 & 5 \\ 0 & 2 & 5 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 5 & 10 & 24 & 6 \\ 7 & 6 & 0 & 2 \\ 0 & 5 & 5 & 4 \end{vmatrix}$$

8. Show that

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ a_1^{n-1} & a_2^{n-2} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{\substack{i,j=1 \\ i>j}}^{n} (a_i - a_j)$$

where the symbol $\prod$ means to form the product of all factors of the type indicated, for the subscripts taking on all values from 1 to n, but always with $i > j$.

9. Show that

$$\begin{vmatrix} 2 & 4 & 10 & 5 \\ 0 & 1 & 4 & 3 \\ 7 & 2 & 5 & 5 \\ 3 & 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 5 & 8 & 12 \\ 7 & 1 & 1 & 2 \\ 15 & 0 & 1 & 8 \end{vmatrix}$$

10. Show that

$$\begin{vmatrix} 2 & 0 & 2 & 1 \\ 3 & 4 & 4 & 0 \\ 6 & 6 & 7 & 6 \\ 8 & 4 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 6 & 4 \\ 1 & 8 & 8 & 0 \\ 1 & 6 & 7 & 8 \\ 1 & 3 & 3 & 5 \end{vmatrix}$$

11.12 Minors: When in a determinant any number of rows and the same number of columns are suppressed, the determinant formed by the remaining elements (maintaining their relative positions) is called a *minor determinant.*

If one row and one column are suppressed, the corresponding minor is called a *first minor.* If two rows and two columns are suppressed, the corresponding minor is called a *second minor*; and so on. The suppressed rows and columns have common elements which form a determinant, and the minor that remains is said to be complementary to this determinant. The minor complementary to the leading element a_1 is called the *leading first minor*; and its leading first minor is called the *leading second minor* of the original determinant.

The upper left-hand corner of a determinant, usually occupied by a_{11}, will be called the leading position.

An element, a_{ij}, in the i-th row and j-th column may be brought into the leading position by $i - 1$ interchanges of adjacent rows and $j - 1$ interchanges of adjacent columns. Since each such interchange changes the sign of the determinant, the resulting determinant will be *positive* or *negative* according as $i + j - 2$, the total number of interchanges, is even or odd; that is, according as $i + j$ is even or odd.

The co-factor of any element a_{ij} of a determinant may be found by bringing that element into the leading position and then suppressing the first row and column. The co-factor so obtained will be positive or negative according as $i + j$ is even or odd. The process of bringing a_{ij} into the leading position does not in any way change the relative position of the elements in the remaining rows and columns. Hence, in order to find the co-factor of an element a_{ij}, suppress the row and column containing a_{ij} and give the resulting determinant the positive or negative sign according as $i + j$ is even or odd.

The co-factor of an element a_{ij} is represented by the symbol A_{ij}. The sign of this co-factor is $(-1)^{i+j}$, but the symbol A_{ij} is generally considered as including the sign and is accordingly written as positive.

The co-factors of the various elements of

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

are as follows:

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}; \quad A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}; \quad A_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A_{21} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}; \quad A_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}; \quad A_{23} = -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}; \quad A_{32} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}; \quad A_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

11.13 Expansion of a determinant according to the elements of a row or column. Since every term of any determinant, Δ, contains one, and only one, element from each row and from each column, it follows that Δ is a linear homogeneous function of the elements of any one row or any one column.

For example, in

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

if we denote the co-factor of any element by the corresponding capital letter, so that A_1 is the co-factor of a_1, B_1 is the co-factor of b_1, etc., we have

$$\begin{aligned} D &= a_1A_1 + b_1B_1 + c_1C_1, & D &= a_1A_1 + a_2A_2 + a_3A_3, \\ D &= a_2A_2 + b_2B_2 + c_2C_2, & D &= b_1B_1 + b_2B_2 + b_3B_3, \\ D &= a_3A_3 + b_3B_3 + c_3C_3, & D &= c_1C_1 + c_2C_2 + c_3C_3. \end{aligned}$$

The algebraic sum of the $(n - 1)!$ terms of Δ which contain the element a_{ij} is $a_{ij}A_{ij}$.

The algebraic sum of all the terms which contain the successive elements

$$a_{i1},\ a_{i2},\ a_{i3},\ \cdots,\ a_{in}$$

of the i-th row, are respectively

$$a_{i1}A_{i1},\ a_{i2}A_{i2},\ a_{i3}A_{i3},\ \cdots,\ a_{in}A_{in}.$$

Each of these n sums contains $(n - 1)!$ terms each of which is a term of the determinant Δ. There are then in all $n(n - 1)! = n!$ different terms. These are the $n!$ terms of the expansion of Δ. Hence

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + a_{i3}A_{i3} + \cdots + a_{in}A_{in}.$$

In a similar manner it may be shown that

$$\Delta = a_{1j}A_{1j} + a_{2j}A_{2j} + a_{3j}A_{3j} + \cdots + a_{nj}A_{nj}.$$

In particular

$$\Delta = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} + \cdots + a_{1n}A_{1n} \text{ and}$$

$$\Delta = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} + \cdots + a_{n1}A_{n1}.$$

11.14 Laplace's development of a determinant.

Definition. If k rows and k columns of a determinant of the n-th order are deleted, the determinant of order $n - k$ formed of the $(n - k)^2$ remaining elements is called a k-th *minor* of the given determinant.

Complementary minors: Any k-th minor of a determinant, and the determinant formed of the k^2 elements common to the k rows and k columns deleted in forming this k-th minor, are called, with respect to each other, *complementary minors* of the given determinant.

Theorem IX: *The product of any two complementary minors, or else the negative of this expression, consists of terms which are found in the development of the original determinant.*

Let the original determinant of order n be represented by Δ. Let the two complementary minors be represented by P_k and P_{n-k} of orders k and $n - k$ respectively. Transpose the rows and columns of Δ which contain the elements of P_k, so that they become in order the first k rows and columns of a new determinant Δ_1. Suppose that this necessitates r interchanges of rows and s interchanges of columns. Then

$$\Delta = (-1)^{r+s}\Delta_1.$$

The determinant containing $(n - k)^2$ elements that occupies the lower right-hand corner of Δ_1 is the minor P_{n-k}.

Develop the product $P_k \cdot P_{n-k}$. Each term in this development will be the product of a term in the expansion of P_k and a term in the expansion of P_{n-k}. But each term in the expansion of P_k contains an element from each of the first k rows and one element from each of the first k columns of Δ_1, while each term in the expansion of P_{n-k} contains an element from each of the last $n - k$ rows and one element from each of the last $n - k$ columns of Δ_1. Thus each term in the product $P_k \cdot P_{n-k}$ contains one element from each of the n rows and one element from each of the n columns of Δ_1 and therefore is a term in the expansion of Δ_1. Hence, save possibly for a change of sign throughout, each term

in the product $P_k \cdot P_{n-k}$ is a term in the expansion of Δ. It follows that the product

$$(-1)^{r+s} P_k \cdot P_{n-k}$$

is composed of terms in the development of Δ. For, in the development of Δ, the only factors by which P_k is multiplied are the terms in the expansion of P_{n-k} and the sign factor $(-1)^{r+s}$. Hence, in the determinant Δ, the minor P_k has the co-factor $(-1)^{r+s} P_{n-k}$.

Theorem X: *If any k rows (or columns) of a determinant of order n be selected, and every possible minor of order k be formed from them, the expansion of the original determinant is obtained by multiplying each of these minors of order k by its co-factor and taking the algebraic sum of these products.*

The given determinant is of order n. Its expansion contains $n!$ terms. The number of different minors of order k which can be formed from the k rows selected is

$$\frac{n!}{k!(n-k)!}.$$

The expansion of each minor contains $k!$ terms; the expansion of its co-factor contains $(n-k)!$ terms. Thus the product of each minor by its co-factor contains $k!(n-k)!$ terms. Then the sum of all the products that can be formed by multiplying each of the different minors by its co-factor contains

$$k!(n-k)! \times \frac{n!}{k!(n-k)!} = n!$$

different terms, all of which are terms in the expansion of the given determinant of order n.

Example.

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_4 & d_4 \end{vmatrix}$$

$$+ \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_3 & d_3 \end{vmatrix}$$

$$+ \begin{vmatrix} a_3 & b_3 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}$$

Exercises

Show that:

1. $\begin{vmatrix} 2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ -1 & 2 & 4 & 0 \\ 0 & 5 & 3 & 1 \end{vmatrix} = 24$

2. $\begin{vmatrix} 1 & 3 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 3 \end{vmatrix} = 3\begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = -15$

3. $\begin{vmatrix} 0 & 0 & 2 & -1 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 8 & -5 \\ 8 & 0 & 8 & -5 & 0 \end{vmatrix} = 0$

4. $\begin{vmatrix} 2 & 7 & 5 & 6 \\ 1 & 1 & 3 & 1 \\ 1 & 5 & 4 & 3 \\ 4 & 7 & 6 & 5 \end{vmatrix} = 45$

5. $\begin{vmatrix} 7 & 2 & 8 & -2 \\ 1 & 4 & -3 & 1 \\ 3 & 0 & 4 & -1 \\ 4 & 6 & -8 & 2 \end{vmatrix} = 14$

6. $\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} = 4$

7. $\begin{vmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{vmatrix} = -3$

8. $\begin{vmatrix} 8 & 7 & 2 & 20 \\ 3 & 1 & 4 & 7 \\ 5 & 0 & 11 & 0 \\ 8 & 1 & 0 & 6 \end{vmatrix} = 2188$

9. $\begin{vmatrix} 0 & x & y & z \\ x & 0 & z & y \\ y & z & 0 & x \\ z & y & x & 0 \end{vmatrix} = x^4 + y^4 + z^4 - 2y^2z^2 - 2z^2x^2 - 2x^2y^2$

10. $\begin{vmatrix} 0 & c & b & d \\ c & 0 & a & e \\ b & a & 0 & f \\ d & e & f & 0 \end{vmatrix} = a^2d^2 + b^2e^2 + c^2f^2 - 2bcef - 2acdf - 2abde$

11. $\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 9 & 4 \\ 1 & 9 & 0 & 1 \\ 1 & 4 & 1 & 0 \end{vmatrix} = 0$

12. $\begin{vmatrix} 1 & 15 & 14 & 4 \\ 12 & 6 & 7 & 9 \\ 8 & 10 & 11 & 5 \\ 13 & 3 & 2 & 16 \end{vmatrix} = 0$

Use Laplace's method to evaluate the following:

13. $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 1 & 0 \\ 3 & 1 & 2 & 1 \\ 0 & 5 & 6 & 0 \end{vmatrix} = -\begin{vmatrix} 0 & 0 & 3 & 1 \\ 0 & 0 & 5 & 6 \\ 1 & 4 & 2 & 3 \\ 3 & 1 & 1 & 2 \end{vmatrix} = 143$

14. $\begin{vmatrix} 4 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 5 & 6 & 4 & 2 \\ 7 & 8 & 1 & 3 \end{vmatrix} = 100$

15. $\begin{vmatrix} 1 & 2 & 0 & 1 & 2 & 3 \\ 4 & 1 & 0 & 4 & 5 & 6 \\ 1 & 3 & 2 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 4 \end{vmatrix} = 42$

16. $\begin{vmatrix} 2 & 6 & 2 & 0 & 0 & 0 \\ 1 & 5 & 3 & 0 & 0 & 0 \\ 3 & 4 & 5 & 0 & 0 & 0 \\ 1 & 1 & 1 & 3 & 2 & 4 \\ 2 & 3 & 4 & 7 & 6 & 1 \\ 5 & 1 & 6 & 5 & 3 & 8 \end{vmatrix} = -84$

17. $\begin{vmatrix} -1 & 0 & 4 & 1 & 3 \\ 2 & 3 & 5 & 2 & 2 \\ 4 & 6 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 2 & 2 \end{vmatrix} = 60$

18. $\begin{vmatrix} 1 & 2 & 3 & 6 & 9 \\ 3 & 1 & 2 & 4 & 5 \\ 2 & 3 & 1 & 5 & 5 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 2 & 1 \end{vmatrix} = 18$

19. $\begin{vmatrix} x & 0 & 0 & a_3 \\ -1 & x & 0 & a_2 \\ 0 & -1 & x & a_1 \\ 0 & 0 & -1 & a_0 \end{vmatrix} = a_0x^3 + a_1x^2 + a_2x + a_3$

20. $\begin{vmatrix} (a_1 - b_1)^2 & (a_1 - b_2)^2 & (a_1 - b_3)^2 \\ (a_2 - b_1)^2 & (a_2 - b_2)^2 & (a_2 - b_3)^2 \\ (a_3 - b_1)^2 & (a_3 - b_2)^2 & (a_3 - b_3)^2 \end{vmatrix} = 2(a_1 - a_2)(a_2 - a_3)(a_3 - a_1)(b_1 - b_2)(b_2 - b_3)(b_3 - b_1).$

11.15 Addition of determinants. *A determinant that has*

$$a_1 + \alpha_1,\ a_2 + \alpha_2,\ a_3 + \alpha_3,\ \cdots,\ a_n + \alpha_n$$

as elements of the first column is equal to the sum of the determinant having $a_1, a_2, a_3, \cdots, a_n$ as the elements of the first column and the determinant having $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n$ as the elements of the first column while the remaining elements are the same in all three determinants.

Let the given determinant be represented by D:

$$D = \begin{vmatrix} a_1 + \alpha_1 & b_1 & c_1 & \cdots \\ a_2 + \alpha_2 & b_2 & c_2 & \cdots \\ a_3 + \alpha_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ a_n + \alpha_n & b_n & c_n & \cdots \end{vmatrix}.$$

Expanding in terms of elements of the first column, we have

$$D = (a_1 + \alpha_1)A_1 + (a_2 + \alpha_2)A_2 + (a_3 + \alpha_3)A_3 + \cdots + (a_n + \alpha_n)A_n$$

$$= (a_1A_1 + a_2A_2 + a_3A_3 + \cdots + a_nA_n) + (\alpha_1A_1 + \alpha_2A_2 + \alpha_3A_3 + \cdots + \alpha_nA_n)$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 & \cdots \\ a_2 & b_2 & c_2 & \cdots \\ a_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & c_n & \cdots \end{vmatrix} + \begin{vmatrix} \alpha_1 & b_1 & c_1 & \cdots \\ \alpha_2 & b_2 & c_2 & \cdots \\ \alpha_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_n & b_n & c_n & \cdots \end{vmatrix}$$

which proves the proposition.

11.16 Addition of columns (or rows): *A determinant is unchanged in value if to the elements of any column (row) are added any arbitrary (but fixed) multiple of the elements of any other given column (row).* For

$$\begin{vmatrix} a_1 + nb_1 & b_1 & c_1 & \cdots \\ a_2 + nb_2 & b_2 & c_2 & \cdots \\ a_3 + nb_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 & \cdots \\ a_2 & b_2 & c_2 & \cdots \\ a_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix} + n \begin{vmatrix} b_1 & b_1 & c_1 & \cdots \\ b_2 & b_2 & c_2 & \cdots \\ b_3 & b_3 & c_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}$$

and the last determinant vanishes, since two columns are alike.

Example.

$$\begin{vmatrix} 1 & 1 & -5 \\ 3 & 2 & -5 \\ 2 & -1 & -4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -5 \\ 2 & 1 & 0 \\ 2 & -1 & -4 \end{vmatrix}$$ by subtracting the first row from the second,

$$= \begin{vmatrix} -1 & 0 & -5 \\ 2 & 1 & 0 \\ 4 & 0 & -4 \end{vmatrix}$$ by adding the second row to the third row and subtracting the second row from the first row.

$$= \begin{vmatrix} -1 & -5 \\ 4 & -4 \end{vmatrix} = 4 + 20 = 24.$$

Exercises

Show that

1. $$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{vmatrix} = 1$$

2. $$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 3 & 4 \\ 1 & 4 & 4 & 5 \end{vmatrix} = 1$$

3. $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix} = 0$

4. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{vmatrix} = 16$

5. $\begin{vmatrix} 5 & 6 & 2 & 3 & 4 \\ 8 & 7 & 3 & 4 & 5 \\ 10 & 5 & 4 & 5 & 6 \\ 12 & 4 & 5 & 6 & 7 \\ 1 & 8 & 6 & 7 & 8 \end{vmatrix} = 0$

6. $\begin{vmatrix} 0 & 1 & 2 & 3 & 4 \\ -1 & 0 & -5 & -4 & -3 \\ -2 & 5 & 0 & 1 & -2 \\ -3 & 4 & -1 & 0 & -1 \\ -4 & 3 & 2 & 1 & 0 \end{vmatrix} = 0$

7. $\begin{vmatrix} 6 & 1 & 1 & -1 \\ 13 & -3 & -1 & -22 \\ 7 & 7 & -4 & -9 \\ 14 & -5 & -8 & -13 \end{vmatrix} = 10209$

8. $\begin{vmatrix} 7 & -2 & 0 & 5 \\ -2 & 6 & -2 & 2 \\ 0 & -2 & 5 & 3 \\ 5 & 2 & 3 & 4 \end{vmatrix} = -972$

9. $\begin{vmatrix} 10 & 18 & 1 & 14 & 22 \\ 4 & 12 & 25 & 8 & 16 \\ 23 & 6 & 19 & 2 & 15 \\ 17 & 5 & 13 & 21 & 9 \\ 11 & 24 & 7 & 20 & 3 \end{vmatrix} = -4680000$

10. $\begin{vmatrix} 8 & 7 & 2 & 20 \\ 3 & 1 & 4 & 7 \\ 5 & 0 & 11 & 0 \\ 8 & 1 & 0 & 6 \end{vmatrix} = 2188$

11. $\begin{vmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix} = -16$

12. $\begin{vmatrix} 4 & 3 & 0 & 0 \\ -1 & 3 & 2 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & -1 & 1 \end{vmatrix} = 53$

13. $$\begin{vmatrix} x & 0 & 0 & 0 & a_4 \\ -1 & x & 0 & 0 & a_3 \\ 0 & -1 & x & 0 & a_2 \\ 0 & 0 & -1 & x & a_1 \\ 0 & 0 & 0 & -1 & a_0 \end{vmatrix} = a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4$$

11.17 n linear equations in n unknowns, D $\neq$ 0. Let there be given the system

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = c_1 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = c_n \end{array} \tag{8}$$

in which the determinant D of the coefficients is not zero.

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0.$$

Then

$$D \cdot x_1 = \begin{vmatrix} a_{11}x_1 & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1}x_1 & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Add to the elements in the first column x_2 times the elements in the second column, $\cdots$, x_i times the elements in the i-th column, $\cdots$, and x_n times the elements in the last column. This gives us

$$D \cdot x_1 = \begin{vmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} c_1 & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ c_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} \text{ because of equations (8).}$$

Whence

$$x_1 = \frac{\begin{vmatrix} c_1 & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ c_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{D} = \frac{C_1}{D} \text{ (this equation defines } C_1\text{)}.$$

In like manner

$$x_2 = C_2/D; \quad \cdots; \quad x_i = C_i/D; \quad \cdots; \quad x_n = C_n/D$$

where C_i is obtained from D by substituting $c_1, \cdots, c_n$ for $a_{1i}, \cdots, a_{ni}$.

That these values of x_i satisfy equations (8) can be established by direct substitution for

$$c_iD - a_{i1}C_1 - a_{i2}C_2 - \cdots - a_{in}C_n$$

may be regarded as the expansion of the following determinant of order $n + 1$, in terms of the elements of its i-th row

$$(-1)^{i-1}\begin{vmatrix} c_1 & a_{11} & a_{12} & \cdots & a_{1n} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_i & a_{i1} & a_{i2} & \cdots & a_{in} \\ c_i & a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_n & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

(by suitably rearranging the columns in $C_2, C_3, \cdots, C_n$), and the determinant vanishes, since two rows are alike. So the solution of (8) is given by

$$x_i = C_i/D \qquad (i = 1, \cdots, n) \qquad (9)$$

Exercises

Solve by determinants the following systems of equations:

1. $x + y + z = 10$
   $2x + y - z = 11$
   $x + 2y - 3z = 5$

2. $x - y + z = 2$
   $x + y - 5z = 4$
   $2x + y - 8z = 6$

3. $x - y - z = 0$
   $x - 2y + z = 6$
   $2x + y - 3z = 2$

4. $x + y + z + w = 5$
   $x - 2y + 2z + w = 4$
   $2x - y + z - 2w = 9$
   $3x + y - 2z + 3w = 3$

5. $x + y + z + w = 8$
   $2x + y - 6z + 6w = 1$
   $x + 2y - 5z + 5w = 1$
   $x - 2y + 3z + w = 1$

6. $x + y + z + w = 4$
   $x - y + z - w = 2$
   $x + 2y + 2z - w = 0$
   $x + y + z = 2$

7. $x + y + z + w = 4$
   $x + y + 2z = 3$
   $y + 2z + w = 2$
   $x - y - w = 3$

8. $x + y + w = 1$
   $y + z = 1$
   $z + w = 4$
   $x + w = 3$

9. $x + y + z = 4$
   $y + z + w = 0$
   $x + z + w = 4$
   $x + y + 2w = 0$

11.18 Rank of a determinant. In a determinant D of order n let us erase $(n - r)$ rows and $(n - r)$ columns; the determinant of order r that remains is an *r-rowed minor* of D. If in a determinant of order n, there is one r-rowed minor, $0 < r < n$, which does not vanish, while every $(r + 1)$-rowed minor does vanish, then D is said to be of *rank* r.

n linear equations in n unknowns, D = 0.

Theorem XI: *Let the determinant D of the coefficients of the unknowns in equations* (8) *be of rank r, $r < n$.*

Consider the determinants S of order $r + 1$ obtained from the $(r + 1)$-rowed minors of D by replacing the elements of any column by the corresponding constants c_i.

(1) *If these determinants S are not all zero, the equations are inconsistent.*

(2) *If all of these determinants S vanish, then the r equations containing the elements of a nonvanishing r- rowed minor of D determine r of the unknowns linearly in terms of the remaining $n - r$ unknowns and these expressions for the r unknowns satisfy the remaining $(n - r)$ equations.*

Proof: One can rearrange the order of the equations (8) and the order of the variables in the equations so that

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdot & \cdot & \cdot \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \text{ is a nonvanishing } r\text{-rowed minor of } D.$$

The elements of this r-rowed minor are taken from the first r of equations (8). Let

$$a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = c_s$$

be any one of the remaining $(n - r)$ equations in (8). Place this equation as the $(r + 1)$th of equations (8). It may then be represented by

$$a_{r+11}\, x_1 + a_{r+12}\, x_2 + \cdots + a_{r+1n}\, x_n = c_{r+1}.$$

Let R represent any $(r + 1)$-rowed determinant whose elements are taken from the following array of elements:

$$\left\| \begin{matrix} a_{11} & \cdots & a_{1r} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \cdots & a_{rr} & \cdots & a_{rn} \\ a_{r+1\,1} & \cdots & a_{r+1\,r} & \cdots & a_{r+1\,n} \end{matrix} \right\| \tag{10}$$

Then $R = 0$, since D is of rank r.

Define S_1 as follows:

$$S_1 = \begin{vmatrix} a_{11} & \cdots & a_{1r} & c_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \cdots & a_{rr} & c_r \\ a_{r+11} & \cdots & a_{r+1r} & c_{r+1} \end{vmatrix}$$

S_1 is an S. We assume $S_1 \neq 0$

Let C_1, $C_2 \cdots$, C_{r+1} be the minors of c_1, c_2, $\cdots$, c_{r+1} in S_1. Multiply the first $(r + 1)$ equations (8) by C_1, $-C_2$, $\cdots$, $(-1)^r C_{r+1}$ respectively and add. We have

$$\begin{aligned} &[a_{11} C_1 - a_{21} C_2 + \cdots + (-1)^r a_{r+11} C_{r+1}]x_1 + \cdots \\ &+ [a_{1j} C_1 - a_{2j} C_2 + \cdots + (-1)^r a_{r+1j} C_{r+1}]x_j + \cdots \\ &+ [a_{1n} C_1 - a_{2n} C_2 + \cdots + (-1)^r a_{r+1n} C_{r+1}]x_n \\ &= c_1 \ C_1 - c_2 \ C_2 + \cdots + (-1)^r c_{r+1} \ C_{r+1} = S_1 \neq 0. \end{aligned}$$

But the coefficient of x_j is

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & a_{1j} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \cdots & a_{rr} & a_{rj} \\ a_{r+1\,1} & \cdots & a_{r+1\,r} & a_{r+1\,j} \end{vmatrix} = 0,$$

being an S if $j > r$, and having two like columns if $j \leq r$.

Hence $0 = S_1$; that is, for consistency we must have $S_1 = 0$. But we have assumed $S_1 \neq 0$. Therefore the equations are inconsistent, and we have proved the first part of the theorem.

Suppose all of the determinants R and S vanish and that $C_{r+1} \neq 0$. Let

$$E_j \equiv a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n - c_j;$$

then

$$C_1E_1 - C_2E_2 + \cdots + (-1)^r C_{r+1}E_{r+1} = \pm S_1.$$

But by the assumption $S_1 = 0$. So E_{r+1} is a linear combination of E_1, $\cdots$, E_r. That is, the $(r + 1)$st equation is a linear combination of the first r equations and so is satisfied by any set of solutions of the first r equations. This $(r + 1)$st equation was any one of the remaining $(n - r)$ equations. Since $C_{r+1} \neq 0$, the first r equations determine x_1, $\cdots$, x_r linearly in terms of the $n - r$ remaining x's by (9).

Exercises

Discuss the following sets of equations:

1.	$3x - 2y + z + 6 = 0$ $2x + 5y - 3z - 2 = 0$ $4x - 9y + 5z + 14 = 0$	Three planes with a common line of intersection. Consistent.
2.	$x - y - z = 2$ $2x - 2y - 2z = 5$ $x + 2y + 3z = 4$	Two parallel planes cut by a third plane. Inconsistent.
3.	$x + y - 2z = 3$ $2x + 2y - 4z = 6$ $3x + 3y - 6z = 9$	Three coincident planes. Consistent.
4.	$x + 2y - z = 3$ $x + 2y - z = 4$ $2x + 4y - 2z = 3$	Three parallel planes. Inconsistent.
5.	$-4x - 5y + z = 7$ $2x + 3z = 10$ $-5y + 7z = 3$	The three planes bound a triangular prism. Inconsistent.
6.	$x + y - z = 1$ $5x - y - z = 6$ $7x + y - 3z = 10$	The three planes bound a triangular prism. Inconsistent.

11.19 Homogeneous linear equations. When $c_1, c_2, \cdots, c_n$ in equations (8) are all zero, the equations are called *homogeneous*. Obviously equations (8) now have the solution $x_1 = x_2 = \cdots = x_n = 0$. From (9) this solution is unique if $D \neq 0$.

If $D = 0$, theorem XI takes the following form:

If the determinant D of the coefficients of n homogeneous linear equations in n unknowns is of rank r, $r < n$, the r equations involving the elements of a nonvanishing r-rowed minor of D determine uniquely r of the unknowns as linear functions of the remaining $n - r$ unknowns, which can be assigned arbitrary values, and these values of these r unknowns satisfy the remaining $n - r$ equations.

Example:

$$\begin{aligned} x - y - z &= 0 \\ 2x - y - 4z &= 0 \\ 3x - 5y + z &= 0 \end{aligned} \qquad D = \begin{vmatrix} 1 & -1 & -1 \\ 2 & -1 & -4 \\ 3 & -5 & 1 \end{vmatrix} = 0. \quad D \text{ is of rank } 2.$$

From the first two equations $x = 3z$; $y = 2z$, and these values substituted in $3x - 5y + z = 0$ satisfy the equation. We can assign to z any arbitrary value.

Exercises

Discuss the following homogeneous systems:

1. $$\begin{aligned} x - y - z &= 0 \\ 2x - 3y + z &= 0 \\ 3x - 2y - 6z &= 0 \end{aligned}$$
Rank 2; $x:y:z = 4:3:1$

2. $$\begin{aligned} 2x - 2y - 3z &= 0 \\ x + 2y - 6z &= 0 \\ x - 2y &= 0 \end{aligned}$$
Rank 2; $x:y:z = 6:3:2$

3. $$\begin{aligned} 2x - 3y + z &= 0 \\ 4x - 6y + 2z &= 0 \\ 10x - 15y + 5z &= 0 \end{aligned}$$
Rank 1; two unknowns arbitrary

4. $$\begin{aligned} x + y - z - w &= 0 \\ 2x + 3y - 2z - 4w &= 0 \\ 3x + 2y - 3z - w &= 0 \\ 2x + 2y - z - 6w &= 0 \end{aligned}$$
Rank 3; $x:y:z:w = 3:2:4:1$

5. $$\begin{aligned} x + y - 3z - 4w &= 0 \\ 2x - 3y - z - 3w &= 0 \\ x - y - z - 2w &= 0 \\ x + 2y - 4z - 5w &= 0 \end{aligned}$$
Rank 2; $x = 2z + 3w$, $y = z + w$

6. $$\begin{aligned} x - y + z - 3w &= 0 \\ x - 2y + 2z - 2w &= 0 \\ x - 2y + z &= 0 \\ y + z - 5w &= 0 \end{aligned}$$
Rank 3; $x:y:z:w = 4:3:2:1$

11.20 Matrices. Definition I: *A system of mn quantities arranged in a rectangular array of m rows and n columns is called a matrix and in particular a matrix of order mn, or an m by n matrix.*

Definition II: *A matrix is said to be of rank r if it contains at least one r-rowed square minor whose determinant is not zero, while the determinant of every square minor of order higher than r, which the matrix contains, is zero.*

Let the coefficients of the unknowns in any set of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= c_1 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad & \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= c_m \end{aligned} \tag{11}$$

be arranged in the order in which they occur in the equations. We obtain a matrix which we denote by A.

$$A \equiv \left\| \begin{matrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{matrix} \right\|$$

The matrix obtained from A by adding the column composed of the constant terms $c_1, \cdots, c_m$ in the above equations is called the augmented matrix and will be denoted by C.

$$C \equiv \left\| \begin{matrix} a_{11} & \cdots & a_{1n} & c_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & c_m \end{matrix} \right\|$$

11.21 m linear equations in n unknowns. We can determine whether the given equations are consistent or not by making use of the following theorem:

Theorem XII: *A necessary and sufficient condition for a system of m linear equations in n unknowns to be consistent is that the matrix of the system have the same rank as the augmented matrix $m > n$.*

If the rank of both matrices is r, the values of $n - r$ of the unknowns may be assigned at pleasure and the others will then be uniquely determined, provided that the matrix of the coefficients of the remaining r unknowns is of rank r.

By Theorem XI part (1) the equations (11) are inconsistent if C is of rank $r + 1$ and A is of rank less than $r + 1$. By the second part of Theorem XI, if A and C are both of rank r, all of m equations are linear combinations of r of them. Note that the rank of C cannot be less than that of A since every determinant in A is in C.

If the equations are consistent and the rank of the matrix A is r , we find a certain r of the unknowns linearly in terms of the remaining $(n - r)$ unknowns. Substitute these values of these r unknowns in the remaining $m - r$ equations and it will be seen that the equations are satisfied.

Corollary. *Any system of $n + 1$ equations in n unknowns is inconsistent if the determinant of order $n + 1$ formed from the augmented matrix is not zero.*

Example 1: The equations

$$\begin{aligned} x + y &= 5 \\ 2x - y &= 4 \\ 2x - 3y &= 1 \end{aligned} \text{ are inconsistent since } \begin{vmatrix} 1 & 1 & 5 \\ 2 & -1 & 4 \\ 2 & -3 & 1 \end{vmatrix} = -3 \neq 0$$

Example 2: Solve the system

$$\begin{aligned} 2x - y + 3z &= 1 \\ 4x - 2y - z &= -3 \\ 2x - y - 4z &= -4 \\ 10x - 5y - 6z &= -10 \end{aligned}$$

The first two equations can be solved for x and z in terms of y. We find

$$x = \frac{7y - 8}{14}, \qquad z = \frac{5}{7}.$$

Substitution of these values of x and z in the two remaining equations shows that they are satisfied. Hence the system is consistent. Any arbitrary value y_1 may be assigned to y. Then the solution is

$$y = y_1; \qquad x = \frac{7y_1 - 8}{14}; \qquad z = \frac{5}{7}.$$

Exercises

Discuss the following systems of equations:

1. $\begin{aligned} x + y + z &= 6 \\ 2x - y - z &= 3 \\ x - y + 2z &= 3 \\ 2x - 2y - z &= 1 \end{aligned}$ The rank of A is 3. The rank of C is 3 $\therefore$ consistent. $x = 3,\ y = 2,\ z = 1$

2. $\begin{aligned} 2x + 3y &= 19 \\ x + y &= 8 \\ x - y &= 1 \end{aligned}$ Augmented matrix $\neq 0$ $\therefore$ inconsistent.

3. $\begin{aligned} x + 2y + z &= 7 \\ 2x + 4y - z &= 8 \\ x + 2y - 2z &= 1 \\ 3x + 6y - 5z &= 5 \end{aligned}$ Rank of $A = 2$. Rank of $C = 2$ $\therefore$ consistent. $x = 5 - 2y;\ z = 2$. y is arbitrary

4. $\begin{aligned} 3x - y + z &= 6 \\ x + 3y + z &= 2 \\ 2x + y + z &= 4 \\ x - 2y &= 1 \end{aligned}$ A of rank 2. C of rank 3. $\therefore$ inconsistent.

5. $\begin{aligned} x + 2y + z + w &= 5 \\ 2x + 4y - 3z - 3w &= 0 \\ 3x + 6y - 4z - 4w &= 1 \\ 2x + 4y - z - w &= 4 \\ x + 2y - z - w &= 1 \end{aligned}$ Rank of $A = 2$. Rank of $C = 2$ $\therefore$ consistent. $x = 3 - 2y$. $z = 2 - w$. y and w are arbitrary.

6. $\begin{aligned} x + 2y - Z &= 3 \\ x - y + Z &= 4 \\ x + y - 2Z &= 0 \\ 2x - y + Z &= 7 \end{aligned}$

7. $\begin{aligned} 2x - y + 2Z &= 5 \\ x + 2y + Z &= 5 \\ x + Z &= 3 \\ 7x - y + 7Z &= 20 \end{aligned}$

Chapter XII

SYMMETRIC FUNCTIONS OF THE ROOTS

12.1 The elementary symmetric functions. Let the n numbers $r_1, r_2, \cdots, r_n$ be the roots of the equation

$$f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0. \ (a_0 > 0) \quad (1)$$

In §5.6 it was shown that

$$\begin{aligned} \sum r_1 &= r_1 + r_2 + \cdots + r_n = -a_1/a_0 \\ \sum r_1r_2 &= r_1r_2 + r_1r_3 + \cdots + r_1r_n + r_2r_3 + \cdots + r_{n-1}r_n = a_2/a_0 \\ \sum r_1r_2r_3 &= r_1r_2r_3 + r_1r_2r_4 + \cdots + r_{n-2}r_{n-1}r_n = -a_3/a_0 \qquad (2) \\ \sum r_1r_2 \cdots r_n &= r_1r_2 \cdots r_n = (-1)^n a_n/a_0. \end{aligned}$$

The symbol $\sum$ is the Greek letter sigma and is used to signify that one is to take the sum of all terms, like the one following the symbol, that can be formed from the given variables by permutations of those variables. A function so formed is a symmetric function of the variables; that is, it is unchanged by the interchange of any two of the variables. The *elementary symmetric functions* are those given by (2) above. As indicated in (2) these elementary symmetric functions are rational functions of the coefficients of (1), and rational integral functions of the remaining coefficients if $a_0 = 1$. There are always exactly n elementary symmetric functions of the roots of an equation properly of the nth degree.

For three variables α, β, γ the elementary symmetric functions are $\alpha + \beta + \gamma$, $\alpha\beta + \beta\gamma + \gamma\alpha$, $\alpha\beta\gamma$.

Following are given some more symmetric functions of the three variables α, β, γ

$$\begin{aligned} \sum \alpha^2 &= \alpha^2 + \beta^2 + \gamma^2. \\ \sum \alpha^2\beta^2 &= \alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2. \\ \sum \alpha^2\beta &= \alpha^2\beta + \alpha^2\gamma + \beta^2\alpha + \beta^2\gamma + \gamma^2\alpha + \gamma^2\beta. \end{aligned}$$

$$\sum \frac{1}{\alpha} = \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}.$$

$$\sum \frac{\alpha}{\beta} = \frac{\alpha}{\beta} + \frac{\beta}{\alpha} + \frac{\alpha}{\gamma} + \frac{\gamma}{\alpha} + \frac{\beta}{\gamma} + \frac{\gamma}{\beta}.$$

It is customary to use the notation $s_1, s_2, \cdots, s_k, \cdots$ to represent $\sum r_1, \sum r_1^2, \cdots, \sum r_1^k, \cdots$ where k may be allowed to range over all positive integral values.

Exercises

1. Find the value of $\sum_6 \alpha^2\beta$* of the roots of the cubic equation

$$x^3 + px^2 + qx + r = 0.$$

Solution. Multiply together the equations

$$\alpha + \beta + \gamma = -p,$$
$$\alpha\beta + \beta\gamma + \gamma\alpha = q.$$

We obtain

$$\sum \alpha^2\beta + 3\alpha\beta\gamma = -pq;$$

hence

$$\sum_6 \alpha^2\beta = 3r - pq.$$

2. If α, β, γ are the roots of $x^3 + px^2 + qx + r = 0$, prove that

(a) $\sum_3 \alpha^2 = p^2 - 2q$

(b) $\sum_3 \alpha^2\beta\gamma = pr$

(c) $\sum_3 \alpha^2\beta^2 = q^2 - 2pr$. Square $\sum_3 \alpha\beta$.

(d) $(\beta + \gamma)(\gamma + \alpha)(\alpha + \beta) = r - pq$

(e) $\sum_3 \alpha^3 = -p^3 + 3pq - 3r$.

Find the product of $\sum \alpha$ and $\sum \alpha^2$

(f) $\sum \alpha^3\beta = p^2q - 2q^2 - pr$.

* The notation $\sum_6$ denotes that there are six terms in the summation.

3. If $\alpha, \beta, \gamma, \delta$ are the roots of $x^4 + px^3 + qx^2 + rx + s = 0$, prove that

(a) $\sum_4 \alpha^2 = p^2 - 2q$

(b) $\sum_4 \alpha^3 = -p^3 + 3pq - 3r$

(c) $\sum_4 \alpha^4 = p^4 - 4p^2q + 2q^2 + 4pr - 4s$

(d) $\sum_6 \alpha^2\beta^2 = q^2 - 2pr + 2s$

(e) $\sum_{12} \alpha^2\beta\gamma = pr - 4s$. Find the product of $\sum_4 \alpha$ and $\sum_4 \alpha\beta\gamma$.

(f) $\sum_{12} \alpha^3\beta = p^2q - 2q^2 - pr + 4s$.

Find the product of $\sum_6 \alpha\beta$ and $\sum_4 \alpha^2$

(g) $\sum_4 \frac{1}{\alpha} = -\frac{r}{s}$

(h) $\sum_6 \frac{1}{\alpha\beta} = \frac{q}{s}$

(i) $\sum_4 \frac{1}{\alpha^2} = \frac{r^2 - 2qs}{s^2}$.

12.2 Newton's theorem on the sums of like powers of the roots. The sums of like powers of the roots of

$$f(x) = x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n = 0 \qquad (3)$$

can be expressed as rational integral functions of the coefficients of $f(x) = 0$.

Let the roots of (3) be $\alpha_1, \alpha_2, \cdots, \alpha_n$. We proceed first to calculate the sums $s_1 = \sum \alpha_1$; $s_2 = \sum \alpha_1^2$; $\cdots$, $s_p = \sum \alpha_1^p$; $\cdots$, $s_{n-1} = \sum \alpha_1^{n-1}$.

By §5.12 we have

$$f'(x) = \frac{f(x)}{x - \alpha_1} + \frac{f(x)}{x - \alpha_2} + \cdots + \frac{f(x)}{x - \alpha_n}.$$

By actual division we find

$$\frac{f(x)}{x - \alpha} = x^{n-1} + (\alpha + p_1)x^{n-2} + (\alpha^2 + p_1\alpha + p_2)x^{n-3} + \cdots$$
$$+ (\alpha^m + p_1\alpha^{m-1} + p_2\alpha^{m-2} + \cdots + p_m)x^{n-m-1} + \cdots$$

In this equation replace α in succession by each of the quantities $\alpha_1, \alpha_2, \cdots, \alpha_n$ and put $s_p = \sum \alpha_1^p = \alpha_1^p + \alpha_2^p + \cdots + \alpha_n^p$. By adding all of these results, we have the following value for $f'(x)$:

$$f'(x) = nx^{n-1} + (s_1 + np_1)x^{n-2} + (s_2 + p_1s_1 + np_2)x^{n-3} + \cdots$$
$$+ (s_m + p_1s_{m-1} + p_2s_{m-2} + \cdots + np_m)x^{n-m-1} + \cdots.$$

By §4.3, we know that

$$f'(x) = nx^{n-1} + (n-1)p_1x^{n-2} +$$
$$(n-2)p_2x^{n-3} + \cdots + 2p_{n-2}\,x + p_{n-1}.$$

Equating coefficients of like powers of x in the two expressions for $f'(x)$, we obtain

$$\begin{aligned} &s_1 + p_1 = 0,\\ &s_2 + p_1s_1 + 2p_2 = 0,\\ &s_3 + p_1s_2 + p_2s_1 + 3p_3 = 0,\\ &s_4 + p_1s_3 + p_2s_2 + p_3s_1 + 4p_4 = 0, \qquad\qquad (4)\\ &\cdots\cdots\cdots\cdots\\ &s_{n-1} + p_1s_{n-2} + p_2s_{n-3} + \cdots + p_{n-2}s_1 + (n-1)p_{n-1} = 0. \end{aligned}$$

The first equation determines s_1 in terms of $p_1, p_2, \cdots, p_n$; the second s_2; the third s_3; and so on, until s_{n-1} is determined. We find in this way

$$s_1 = -p_1; \qquad s_2 = p_1^2 - 2p_2; \qquad s_3 = -p_1^3 + 3p_1p_2 - 3p_3;$$
$$s_4 = p_1^4 - 4p_1^2p_2 + 4p_1p_3 + 2p_2^2 - 4p_4;$$
$$s_5 = -p_1^5 + 5p_1^3p_2 - 5p_1^2p_3 - 5(p_2^2 - p_4)p_1 + 5(p_2p_3 - p_5);$$

and so on, up to s_{n-1}.

Having shown how $s_1, s_2, \cdots, s_{n-1}$ can be calculated in terms of the coefficients, we proceed to extend our results to the sums of all positive powers of the roots, viz. $s_n, s_{n+1}, \cdots$.

To this end multiply $f(x)$ by x^{m-n}, where $m \geq n$. We have

$$x^{m-n}f(x) = x^m + p_1x^{m-1} + p_2x^{m-2} + \cdots + p_nx^{m-n} = 0.$$

In this equation, replace x in succession by $\alpha_1, \alpha_2, \cdots, \alpha_n$ and add the resulting n equations. We have

$$s_m + p_1s_{m-1} + p_2s_{m-2} + \cdots + p_ns_{m-n} = 0.$$

If in this equation we give m the values n, $n + 1$, $n + 2$, $\cdots$ in succession and observe that $s_0 = n$, we obtain

$$\begin{aligned} s_n + p_1 s_{n-1} + p_2 s_{n-2} + \cdots + n p_n &= 0, \\ s_{n+1} + p_1 s_n + p_2 s_{n-1} + \cdots + p_n s_1 &= 0, \\ s_{n+2} + p_1 s_{n+1} + p_2 s_n + \cdots + p_n s_2 &= 0, \text{ etc.} \end{aligned}$$

These last equations are all included in the formula

$$s_k + p_1 s_{k-1} + p_2 s_{k-2} + \cdots + p_n s_{k-n} = 0 \qquad (k \geq n). \tag{5}$$

So the sums of all positive integral powers of the roots may be expressed rationally in terms of the coefficients.

By transforming the equation (for $p_n \neq 0$), by the substitution $x = 1/y$, into one whose roots are the reciprocals of $\alpha_1, \alpha_2, \cdots, \alpha_n$, and applying the above formulas to the roots of the transformed equation, we can find the sums of *negative* powers of the roots of the original equation $f(x) = 0$.

From equations (4) and (5) it is clear that the s_i can be expressed as rational integral functions, with integral coefficients, of the coefficients of $f(x) = 0$.

Exercises

1. For $x^4 + x^3 + x^2 + x + 1 = 0$ show that $s_1 = s_2 = s_3 = s_4 = -1$.
2. For $x^5 - 8 = 0$ show that $s_1 = s_2 = s_3 = s_4 = 0$, $s_5 = 40$.
3. For $x^6 - x - 1 = 0$ show that $s_1 = s_2 = s_3 = s_4 = 0$, $s_5 = 5$; $s_6 = 6$; $s_7 = s_8 = s_9 = 0$; $s_{10} = 5$; $s_{11} = 11$; $s_{12} = 6$.
4. For $x^n + p_3 x^{n-3} + p_4 x^{n-4} + \cdots + p_n = 0$ show that $s_1 = s_2 = 0$, $s_3 = -3p_3$, $s_4 = -4p_4$, $s_5 = -5p_5$.
5. For $x^5 + p_4 x^4 + p_5 = 0$ show that $s_1 = s_2 = s_3 = 0$; $s_4 = -4p_4$; $s_5 = -5p_5$.
6. Find s_m for $x^3 - 3x^2 + 3x - 1 = 0$. Ans. $s_m = 3$ for every integral value of m.
7. Find s_m for $x^3 + 3x^2 + 3x + 1 = 0$. Ans. $s_{2k} = 3$; $s_{2k+1} = -3$ for every integral value of k.
8. For $(x - 1)^n = 0$ show that $s_k = n$.
9. For $(x + 1)^n = 0$ show that $s_{2k} = n$; $s_{2k+1} = -n$.

10. For $x^3 - 2x^2 + x + 3 = 0$ show that $s_1 = 2$; $s_2 = 2$; $s_3 = -7$; $s_4 = -22$; $s_{-1} = -\frac{1}{3}$; $s_{-2} = 13/9$.

11. For $x^4 - 2x^3 + x^2 + 3x + 1 = 0$ show that $s_1 = s_2 = 2$; $s_3 = -7$; $s_4 = -26$; $s_{-1} = -3$; $s_{-2} = 7$; $s_{-3} = -12$; $s_{-4} = 19$.

12. For $x^n - 1 = 0$ show that $s_m = n$ or 0, according as m is or is not divisible by n.

13. For $x^n + 1 = 0$ show that $s_m = \pm n$ or 0, according as m is or is not divisible by n.

14. For $(x - 1)(x - 2) \cdots (x - n) = 0$ show that $s_1 = \frac{1}{2}n(n + 1)$; $s_2 = \frac{1}{6}n(n + 1)(2n + 1)$; $s_3 = s_1^2$.

12.3 Fundamental theorem on symmetric functions. *Every rational integral symmetric function of the roots of an equation*

$$f(x) = x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n = 0$$

can be expressed as a rational integral function of the coefficients.

Let us first find the value of the homogeneous symmetric function $\sum \alpha_1^p\alpha_2^q$. We have

$$s_p = \alpha_1^p + \alpha_2^p + \alpha_3^p + \cdots + \alpha_n^p,$$

$$s_q = \alpha_1^q + \alpha_2^q + \alpha_3^q + \cdots + \alpha_n^q,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the roots of $f(x) = 0$.

Multiplying, we have for $p \neq q$,

$$s_ps_q = \alpha_1^{p+q} + \alpha_2^{p+q} + \cdots + \alpha_n^{p+q} + \alpha_1^p\alpha_2^q + \alpha_1^q\alpha_2^p + \cdots,$$

or

$$s_ps_q = s_{p+q} + \sum \alpha_1^p\alpha_2^q \qquad (p \neq q).$$

Therefore,

$$\sum \alpha_1^p\alpha_2^q = s_ps_q - s_{p+q} \qquad (p \neq q). \quad (6)$$

This formula expresses the double symmetric function in terms of the single symmetric functions s_p, s_q, s_{p+q}.

If $p = q$, then $\alpha_1^p\alpha_2^q = \alpha_1^q\alpha_2^p$ and $\sum \alpha_1^p\alpha_2^q = 2\sum \alpha_1^p\alpha_2^p$ whence

$$\sum \alpha_1^p\alpha_2^p = \tfrac{1}{2}(s_p^2 - s_{2p}).$$

In either case, the homogeneous symmetric function is expressed as a rational integral function of the s_i. And by Newton's

theorem the s_i can be expressed as rational integral functions of the coefficients.

In each term of the triple symmetric function $\sum \alpha_1^p\alpha_2^q\alpha_3^r$ three roots are involved. We now proceed to show how this triple symmetric function can be expressed as a rational integral function of the s_i and hence as a rational integral function of the coefficients.

Multiply $\sum \alpha_1^p\alpha_2^q$ by s_r, where

$$\sum \alpha_1^p\alpha_2^q = \alpha_1^p\alpha_2^q + \alpha_1^q\alpha_2^p + \alpha_1^p\alpha_3^q + \cdots \; (p \neq q)$$

$$s_r = \alpha_1^r + \alpha_2^r + \alpha_3^r + \cdots + \alpha_n^r.$$

We have for p, q, r, distinct,

$$s_r \sum \alpha_1^p\alpha_2^q = \sum \alpha_1^{p+r}\alpha_2^q + \sum \alpha_1^p\alpha_2^{q+r} + \sum \alpha_1^p\alpha_2^q\,\alpha_3^r,$$

a formula connecting double and triple symmetric functions. But, by (6),

$$\sum \alpha_1^{p+r}\alpha_2^q = s_{p+r}s_q - s_{p+q+r}$$

$$\sum \alpha_1^p\alpha_2^{q+r} = s_p s_{q+r} - s_{p+q+r}$$

$$\sum \alpha_1^p\alpha_2^q = s_p s_q \quad - s_{p+q}.$$

Substituting these values, we find (for p, q, r, distinct),

$$\sum \alpha_1^p\alpha_2^q\,\alpha_3^r = s_p s_q s_r - s_{p+q}s_r - s_{p+r}s_q - s_p s_{q+r} + 2s_{p+q+r}.$$

If $p = q = r$, we have $\sum \alpha_1^p\alpha_2^p\alpha_3^p = \frac{1}{6}(s_p^3 - 3s_p s_{2p} + 2s_{3p})$.

We now proceed to show that if any homogeneous symmetric function, in which each term involves m roots, can be expressed as a rational integral function of the s_i, then any homogeneous symmetric function in which each term involves $m + 1$ roots can also be expressed as a rational integral function of the s_i and hence as a rational integral function of the coefficients.

In each term of the symmetric function $\sum \alpha_1^p\alpha_2^q \cdots \alpha_m^r\alpha_{m+1}^t$ $m + 1$ roots are involved. We now proceed to show how this symmetric function can be expressed in terms of symmetric functions in which m roots are involved.

Multiply $\sum \alpha_1^p\alpha_2^q \cdots \alpha_m^r$ by s_t, where ($p, q, \cdots, r$, distinct)

$$\sum \alpha_1^p\alpha_2^q \cdots \alpha_m^r = \alpha_1^p\alpha_2^q \cdots \alpha_m^r + \alpha_1^q\alpha_2^p \cdots \alpha_m^r + \alpha_1^r\alpha_2^q \cdots \alpha_m^p$$
$$+ \alpha_1^p\alpha_2^r \cdots \alpha_m^q + \cdots$$

$$s_t = \alpha_1^t + \alpha_2^t + \cdots + \alpha_n^t \; (t \neq p, q, \cdots, r).$$

We have

$$s_t \sum \alpha_1^p \alpha_2^q \cdots \alpha_m^r = \sum \alpha_1^{p+t} \alpha_2^q \cdots \alpha_m^r + \sum \alpha_1^p \alpha_2^{q+t} \cdots \alpha_m^r$$
$$+ \cdots + \sum \alpha_1^p \alpha_2^q \cdots \alpha_m^{r+t} + \sum \alpha_1^p \alpha_2^q \cdots \alpha_m^r \alpha_{m+1}^t.$$

Those cases in which some of the letters $p, q, \cdots, r$ are equal (but not all) can be handled separately with no essentially new complications. This equation shows that $\sum \alpha_1^p \alpha_2^q \cdots \alpha_m^r \alpha_{m+1}^t$ can be expressed in terms of homogeneous symmetric functions in which each term involves m roots, plus one such symmetric function multiplied by s_t; hence as a rational integral function of the s_i, and therefore, finally, as a rational integral function of the coefficients.

If any rational integral symmetric function is not homogeneous, then it is the sum of two or more rational integral symmetric functions which are homogeneous, such as

$$a(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) + b(\alpha_1 + \alpha_2 + \alpha_3).$$

Hence it is clear that any rational integral symmetric function can be expressed as a rational integral function of the coefficients of $f(x) = 0$.

12.4 Computation of symmetric functions. Any rational integral symmetric function can be calculated by the method used in establishing the fundamental theorem. In practice, however, it is sometimes easier to use other methods. If the symmetric function contains a large number of roots with small exponents, the indications are usually plain as to which auxiliary simpler symmetric functions should be multiplied together to produce the desired symmetric function along with simpler symmetric functions.

Example 1. Calculate $\sum \alpha_1^2\alpha_2\alpha_3$ of the roots of the general equation

$$x^n + p_1x^{n-1} + p_2x^{n-2} + \cdots + p_{n-1}x + p_n = 0.$$

Multiply together the equations

$$\sum \alpha_1 = -p_1$$
$$\sum \alpha_1\alpha_2\alpha_3 = -p_3.$$

In the product the term $\alpha_1^2\alpha_2\alpha_3$ occurs only once; the term $\alpha_1\alpha_2\alpha_3\alpha_4$ occurs four times, arising from the product of α_1 by $\alpha_2\alpha_3\alpha_4$, of α_2 by $\alpha_1\alpha_3\alpha_4$, of α_3 by $\alpha_1\alpha_2\alpha_4$, and of α_4 by $\alpha_1\alpha_2\alpha_3$. Hence

$$\sum \alpha_1 \sum \alpha_1\alpha_2\alpha_3 = \sum \alpha_1^2\alpha_2\alpha_3 + 4\sum \alpha_1\alpha_2\alpha_3\alpha_4 = p_1p_3;$$

therefore

$$\sum \alpha_1^2\alpha_2\alpha_3 = p_1p_3 - 4p_4.$$

Example 2. Calculate $\sum \alpha_1^3\alpha_2$ for the general equation. Multiply together

$$\sum \alpha_1^2 = p_1^2 - 2p_2$$

$$\sum \alpha_1\alpha_2 = p_2.$$

We have

$$\sum \alpha_1^2 \sum \alpha_1\alpha_2 = \sum \alpha_1^3\alpha_2 + \sum \alpha_1^2\alpha_2\alpha_3 = (p_1^2 - 2p_2)p_2.$$

But from example 1, we have

$$\sum \alpha_1^2\alpha_2\alpha_3 = p_1p_3 - 4p_4,$$

hence

$$\sum \alpha_1^3\alpha_2 = \sum \alpha_1^2 \sum \alpha_1\alpha_2 - \sum \alpha_1^2\alpha_2\alpha_3$$
$$= p_1^2p_2 - 2p_2^2 - p_1p_3 + 4p_4.$$

Example 3. Calculate $\sum \alpha_1^2\alpha_2^2$ for the general equation. Squaring

$$\sum \alpha_1\alpha_2 = p_2$$

we have

$$p_2^2 = \sum \alpha_1^2\alpha_2^2 + 2\sum \alpha_1^2\alpha_2\alpha_3 + 6\sum \alpha_1\alpha_2\alpha_3\alpha_4.$$

In squaring, the term $\alpha_1\alpha_2\alpha_3\alpha_4$ will arise from the product of $\alpha_1\alpha_2$ by $\alpha_3\alpha_4$, of $\alpha_1\alpha_3$ by $\alpha_2\alpha_4$, and of $\alpha_1\alpha_4$ by $\alpha_2\alpha_3$; hence the coefficient of $\alpha_1\alpha_2\alpha_3\alpha_4$ is 6, since each product occurs twice in squaring. Then

$$\sum \alpha_1^2\alpha_2^2 = p_2^2 - 2\sum \alpha_1^2\alpha_2\alpha_3 - 6\sum \alpha_1\alpha_2\alpha_3\alpha_4$$
$$= p_2^2 - 2(p_1p_3 - 4p_4) - 6p_4$$
$$= p_2^2 - 2p_1p_3 + 2p_4.$$

Exercises

For the general equation show that:

1. $\sum \alpha_1^2\alpha_2^2\alpha_3 = -p_2p_3 + 3p_1p_4 - 5p_5$
2. $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 = -p_1p_4 + 5p_5$
3. $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4\alpha_5 = p_1p_5 - 6p_6$
4. $\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4 = p_2p_4 - 4p_1p_5 + 9p_6$
5. $\sum \alpha_1^2\alpha_2^2\alpha_3^2 = p_3^2 - 2p_2p_4 + 2p_1p_5 - 2p_6$
6. $\sum \alpha_1^3\alpha_2 = p_1^2p_2 - 2p_2^2 - p_1p_3 + 4p_4$
7. Show that for the cubic $x^3 + p_1x^2 + p_2x + p_3 = 0$, we have

(a) $\sum \alpha_1^2\alpha_2^2\alpha_3 = -p_2p_3$
(b) $\sum \alpha_1^2\alpha_2\alpha_3 = p_1p_3$
(c) $\sum \alpha_1^2\alpha_2^2\alpha_3^2 = p_3^2$
(d) $\sum \alpha_1^2\alpha_2^2 = p_2^2 - 2p_1p_3$
(e) $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 = 0$
(f) $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4\alpha_5 = 0$
(g) $\sum \alpha_1^3\alpha_2 = p_1^2p_2 - 2p_2^2 - p_1p_3$

8. Show that for the quartic $x^4 + p_1x^3 + p_2x^2 + p_3x + p_4 = 0$, we have

(a) $\sum \alpha_1^2\alpha_2\alpha_3 = p_1p_3 - 4p_4$
(b) $\sum \alpha_1^2\alpha_2^2\alpha_3 = -p_2p_3 + 3p_1p_4$
(c) $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 = -p_1p_4$
(d) $\sum \alpha_1^2\alpha_2\alpha_3\alpha_4\alpha_5 = 0$
(e) $\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4 = p_2p_4$
(f) $\sum \alpha_1^2\alpha_2^2\alpha_3^2 = p_3^2 - 2p_2p_4$
(g) $\sum \alpha_1^3\alpha_2^3 = p_2^3 - 3p_1p_2p_3 + 3p_3^2 + 3p_1^2p_4 - 3p_2p_4$
(h) $\sum \alpha_1^3\alpha_2^2\alpha_3 = p_1p_2p_3 - 3p_3^2 - 3p_1^2p_4 + 4p_2p_4$

9. If $\alpha_1, \alpha_2, \alpha_3$ are roots of the cubic $x^3 + p_1x^2 + p_2x + p_3 = 0$, show that the cubic with the roots

(a) $\alpha_1^2, \alpha_2^2, \alpha_3^2$ is $x^3 - (p_1^2 - 2p_2)x^2 + (p_2^2 - 2p_1p_3)x - p_3^2 = 0$
(b) $\alpha_1\alpha_2, \alpha_1\alpha_3, \alpha_2\alpha_3$ is $x^3 - p_2x^2 + p_1p_3x - p_3^2 = 0$
(c) $\dfrac{3}{\alpha_1}, \dfrac{3}{\alpha_2}, \dfrac{3}{\alpha_3}$ is $p_3x^3 + 3p_2x^2 + 9p_1x + 27 = 0$
(d) $\alpha_1 + \alpha_2, \alpha_1 + \alpha_3, \alpha_2 + \alpha_3$ is $x^3 + 2p_1x^2 + (p_1^2 + p_2)x + (p_1p_2 - p_3) = 0$
(e) $\alpha_1^3, \alpha_2^3, \alpha_3^3$ is $x^3 + (p_1^3 - 3p_1p_2 + 3p_3)x^2 + (p_2^3 - 3p_1p_2p_3 + 3p_3^2)x + p_3^3 = 0$

Chapter XIII

RESULTANTS: DISCRIMINANTS: ELIMINANTS

13.1 Resultant of two polynomials. Let there be given two polynomials of degrees m and n respectively:

$$f(x) = a_0x^m + a_1x^{m-1} + a_2x^{m-2} + \cdots + a_m = 0 \qquad (a_0 \neq 0),$$

$$g(x) = b_0x^n + b_1x^{n-1} + b_2x^{n-2} + \cdots + b_n = 0 \qquad (b_0 \neq 0).$$

We desire to find the condition that these equations should have a common root.

Let the roots of $f(x) = 0$ be $\alpha_1, \alpha_2, \cdots, \alpha_m$ whence

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m) \qquad (a_0 \neq 0).$$

Let the roots of $g(x) = 0$ be $\beta_1, \beta_2, \cdots, \beta_n$ whence

$$g(x) = b_0(x - \beta_1)(x - \beta_2) \cdots (x - \beta_n) \qquad (b_0 \neq 0).$$

If the given equations have a common root, it is *necessary* and *sufficient* that some one of the quantities

$$g(\alpha_1), g(\alpha_2), \cdots, g(\alpha_m)$$

should be zero, or, in other words, that the product

$$g(\alpha_1)\cdot g(\alpha_2) \cdots g(\alpha_m)$$

should vanish. Since $\beta_1, \beta_2, \cdots, \beta_n$ are the roots of $g(x) = 0$, we have

$$\begin{aligned} g(\alpha_1) &= b_0(\alpha_1 - \beta_1)(\alpha_1 - \beta_2) \cdots (\alpha_1 - \beta_n) \\ g(\alpha_2) &= b_0(\alpha_2 - \beta_1)(\alpha_2 - \beta_2) \cdots (\alpha_2 - \beta_n) \\ & \\ g(\alpha_m) &= b_0(\alpha_m - \beta_1)(\alpha_m - \beta_2) \cdots (\alpha_m - \beta_n). \end{aligned}$$

If we change the signs of the mn factors and multiply these equations, taking together the factors which are situated in the same column, we find

$$R = a_0^n g(\alpha_1)g(\alpha_2) \cdots g(\alpha_m) = (-1)^{mn}b_0^m f(\beta_1)f(\beta_2) \cdots f(\beta_n).$$

Hence a *necessary* and *sufficient* condition that the two equations $f(x) = 0$, $g(x) = 0$ have a common root is $R = 0$. We call R the *resultant* of the two equations.

It is readily seen that

$$\pm R = a_0^n b_0^m \prod (\alpha_r - \beta_s),$$

where $\prod$ signifies the continued product of the mn terms of the form $\alpha_r - \beta_s$.

Example: Find the resultant of $ax + b = 0$ and $x^2 - 1 = 0$.

The roots of $x^2 - 1 = 0$ are $+1, -1$.

Substitute $+1$ in $ax + b$. We have $+a + b$.

Substitute -1 in $ax + b$. We have $-a + b$.

From the definition of a resultant, the resultant in this case is $(a + b)(-a + b) = b^2 - a^2$.

Exercises

1. Show that the resultant of $ax^2 + bx + c = 0$ and $x^2 - 1 = 0$ is $(a + b + c)(a - b + c)$.
2. Show that the resultant of $ax^3 + bx^2 + cx + d = 0$ and $x^2 - 1 = 0$ is $(a + b + c + d)(-a + b - c + d)$.
3. Show that the resultant of $ax + b = 0$ and $x^4 - 1 = 0$ is $(b + a)(b - a)(b + ai)(b - ai) = b^4 - a^4$.
4. Show that the resultant of $ax + b = 0$ and $x^3 - 1 = 0$ is $-(a + b)(a\omega + b)(a\omega^2 + b)$.
5. Show that the resultant of $ax^2 + bx + c = 0$ and $x^3 - 1 = 0$ is $(a + b + c)(a\omega^2 + b\omega + c)(a\omega + b\omega^2 + c)$.
6. Show that the resultant of $ax^3 + bx^2 + cx + d = 0$ and $x^3 - 1 = 0$ is $-(a + b + c + d)(a + b\omega^2 + c\omega + d)$ $(a + b\omega + c\omega^2 + d)$.
7. Show that the resultant of $ax^2 + bx + c = 0$ and $x^4 - 1 = 0$ is $(a + b + c)(a - b + c)(-a + bi + c)(-a - bi + c)$.
8. Show that the resultant of $ax^3 + bx^2 + cx + d = 0$ and $x^4 - 1 = 0$ is $(a + b + c + d)(-a + b - c + d)$ $(-ai - b + ci + d)(ai - b - ci + d)$.
9. Show that the resultant of $ax^2 + bx + c = 0$ and $x^4 + 4 = 0$ is $[b + c + i(2a + b)][b + c - i(2a + b)][c - b + i(2a - b)][c - b - i(2a - b)]$.
10. Show that the resultant of $x^2 - 7x + 12 = 0$ and $(x - a)$ $(x - b) = 0$ is $(a^2 - 7a + 12)(b^2 - 7b + 12)$.

13.2 Sylvester's dialytic method of elimination. Let there be given the two equations

$$f(x) = a_0x^m + a_1x^{m-1} + a_2x^{m-2} + \cdots + a_m = 0 \qquad (a_0 \neq 0),$$
$$g(x) = b_0x^n + b_1x^{n-1} + b_2x^{n-2} + \cdots + b_n = 0 \qquad (b_0 \neq 0).$$

Multiply the first equation, in turn, by

$$x^{n-1},\ x^{n-2},\ \cdots,\ x^2,\ x,\ 1;$$

and the second by $x^{m-1}, x^{m-2}, \cdots, x^2, x, 1.$
We thus obtain the $m + n$ equations

$$x^{n-1}f(x) = 0, \cdots, x^2f(x) = 0, \qquad xf(x) = 0, \qquad f(x) = 0,$$
$$x^{m-1}g(x) = 0, \cdots, x^2g(x) = 0, \qquad xg(x) = 0, \qquad g(x) = 0,$$

which are linear and homogeneous in the $m + n$ quantities,

$$x^{m+n-1}, x^{m+n-2}, \cdots, x^2, x, 1.$$

If $f(x) = 0$, $g(x) = 0$ have a common root it will satisfy all of these $m + n$ equations. Let the different powers of x be taken as $m + n - 1$ unknowns. A *necessary* condition that these $m + n$ linear homogeneous equations in these $m + n$ quantities shall have a solution other than the trivial one in which each unknown is zero is that the following determinant, D, be zero:

$$D = \begin{vmatrix} a_0 & a_1 & a_2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & a_{m-1} & a_m & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & a_{m-2} & a_{m-1} & a_m & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & a_0 & a_1 & a_2 & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & a_m \\ b_0 & b_1 & b_2 & \cdots & b_{n-1} & b_n & 0 & \cdots & 0 & 0 & 0 & \cdots & & 0 \\ 0 & b_0 & b_1 & \cdots & b_{n-2} & b_{n-1} & b_n & \cdots & 0 & 0 & 0 & \cdots & & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & b_0 & b_1 & b_2 & \cdots & & b_n \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ rows} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \text{ rows} \end{matrix} \quad (1)$$

For example, if $f(x)$ is $x^3 - 3x^2 + 4$ and $g(x)$ is $5x^2 - 7x - 6$ then

$$D = \begin{vmatrix} 1 & -3 & 0 & 4 & 0 \\ 0 & 1 & -3 & 0 & 4 \\ 5 & -7 & -6 & 0 & 0 \\ 0 & 5 & -7 & -6 & 0 \\ 0 & 0 & 5 & -7 & -6 \end{vmatrix} = 0,$$

and $x = 2$ is a common root.

We will now show that if $D = 0$, the two equations do have a root in common. To do this it is sufficient to show that $D = R$.

In D replace b_n by $b_n - z$ and consider the equation

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdot & \cdot & \cdot & \cdot & a_{m-1} & a_m & 0 \cdots & 0 \\ 0 & a_0 & a_1 & \cdot & \cdot & \cdot & \cdot & a_{m-2} & a_{m-1} & a_m \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & a_0 & a_1 & a_2 & \cdot & \cdot & \cdot & \cdot & a_m \\ b_0 & b_1 & b_2 \cdots & b_{n-1} & b_n - z & 0 & \cdots & 0 & 0 & 0 \cdots & 0 \\ 0 & b_0 & b_1 \cdots & b_{n-2} & b_{n-1} & b_n - z & \cdots & 0 & 0 & 0 \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & 0 & 0 & 0 & \cdots & b_0 & b_1 & b_n \cdots & b_n - z \end{vmatrix} \tag{2}$$

$$= 0$$

Equation (2) reduces to (1) for $z = 0$; and (2) is of the form

$$a_0^n z^m + \cdots + k_i z^{m-i} + \cdots + (-1)^m D = 0 \qquad (i = 1, 2, \cdots, m - 1). \tag{3}$$

Equation (2), or (3), has $g(\alpha_1)$ for a root. This is seen as follows: in (2) put $z = g(\alpha_1)$. Beginning with the left-hand column, multiply the columns in turn by $\alpha_1^{m+n-1}, \alpha_1^{m+n-2}, \cdots, \alpha_1^2, \alpha_1$ and add to the last column. We find that the elements of the new last column are zero. Thus $g(\alpha_1)$ is a root of (2), and so of (3). In like manner we can show that $g(\alpha_i)$ $(i = 2, \cdots, m)$ are roots of (2), and so of (3). But $(-1)^m D/a_0^n$ is, for equation (3), equal to the product of the roots with their signs changed. Thus whether m is odd or even, we have

$$D = a_0^n g(\alpha_1) g(\alpha_2) \cdots g(\alpha_m) = R.$$

Example: Find the resultant of

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0,$$
$$b_0x^2 + b_1x + b_2 = 0.$$

Multiplying the first of these equations by x, and the second by x^2 and x successively, we have the following five equations:

$$\begin{aligned} a_0x^4 + a_1x^3 + a_2x^2 + a_3x \phantom{{}+ a_3} &= 0; \\ a_0x^3 + a_1x^2 + a_2x + a_3 &= 0; \\ b_0x^4 + b_1x^3 + b_2x^2 \phantom{{}+ b_2} &= 0; \\ b_0x^3 + b_1x^2 + b_2x &= 0; \\ b_0x^2 + b_1x + b_2 &= 0. \end{aligned}$$

These five equations in the four variables x, x^2, x^3, x^4 are consistent if

$$\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix} = 0.$$

This is the required resultant.

13.3 Elimination. If the two equations

$$a_0x + a_1 = 0, \qquad b_0x + b_1 = 0 \qquad\qquad (a_0b_0 \neq 0),$$

are satisfied by the same value of x, then

$$x = -\frac{a_1}{a_0} = -\frac{b_1}{b_0}, \quad \text{from which} \quad E = a_0b_1 - a_1b_0 = 0.$$

E is called an *eliminant* of the two equations.

The result of eliminating x between two equations, when put in a rational integral form is called an *eliminant.*

The result of eliminating x between the two given equations might just as well have been written $a_1b_0 - a_0b_1 = 0$.

The resultant R of the two equations is

$$R = \begin{vmatrix} a_0a_1 \\ b_0b_1 \end{vmatrix} = a_0b_1 - a_1b_0.$$

The result of an elimination, in other cases, may give R multiplied by some extraneous factor which may be either a constant or a function of the coefficients.

Example 1. Find the condition that the two equations $a_0x + a_1 = 0$, $b_0x^2 + b_1x + b_2 = 0$ shall have a root in common.

The only root of $a_0x + a_1 = 0$ is $x = -a_1/a_0$.

If the second equation is to have this same root, then the second equation must be satisfied when in place of x we put $-a_1/a_0$. Making this substitution, we have

$$b_0\frac{a_1^2}{a_0^2} - b_1\frac{a_1}{a_0} + b_2 = 0,$$

whence

$$E = b_0a_1^2 - a_0a_1b_1 + a_0^2b_2.$$

In this case $E = R$, since $R = \begin{vmatrix} a_0 & a_1 & 0 \\ 0 & a_0 & a_1 \\ b_0 & b_1 & b_2 \end{vmatrix}$.

Exercises

1. Show that an eliminant of $a_0x^2 + a_1x + a_2 = 0$ and $b_0x + b_1 = 0$ is $a_0b_1^2 - a_1b_0b_1 + a_2b_0^2$.
2. Show that an eliminant of $a_0x + a_1 = 0$ and $b_0x^3 + b_1x^2 + b_2x + b_3 = 0$ is $a_0^3b_3 - a_0^2a_1b_2 + a_0a_1^2b_1 - a_1^3b_0$ and show that this is equal to the resultant.
3. Show that the resultant of $b_0x^3 + b_1x^2 + b_2x + b_3 = 0$ and $a_0x + a_1 = 0$ is the negative of the resultant of $a_0x + a_1 = 0$ and $b_0x^3 + b_1x^2 + b_2x + b_3 = 0$.
4. Show that the resultant of $a_0x^2 + a_1x + a_2 = 0$ and $b_0x^2 + b_1x + b_2 = 0$ is $a_0^2b_2^2 - a_0a_1b_1b_2 + a_0a_2(b_1^2 - 2b_0b_2) + a_1^2b_0b_2 - a_1a_2b_0b_1 + a_2^2b_0^2$.
5. Show that the resultant of $a_0x^2 + a_1x + a_2 = 0$ and $b_0x^2 + b_1x + b_2 = 0$ is

$$R = \begin{vmatrix} a_0 & a_1 & a_2 & 0 \\ 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & 0 \\ 0 & b_0 & b_1 & b_2 \end{vmatrix} = (a_0b_2 - a_2b_0)^2 - (a_0b_1 - a_1b_0)(a_1b_2 - a_2b_1).$$

6. Obtain an eliminant of $a_0x^2 + a_1x + a_2 = 0$ and $b_0x^2 + b_1x + b_2 = 0$ by equating the roots and rationalizing:

$$x = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_0} = \frac{-b_1 + \sqrt{b_1^2 - 4b_0b_2}}{2b_0}.$$

7. Show that the resultant of $x^n = 0$ and $a_0x^m + a_1x^{m-1} + \cdots + a_m = 0$ is a_m^n.
   (a) Use the definition in article 13.1.
   (b) Use Sylvester's method.
8. By Sylvester's method show that the following pairs of equations have a root in common:
   (a) $x^2 - 7x + 12 = 0;\ x^2 + x - 12 = 0$
   (b) $x^3 - 6x^2 + 11x - 6 = 0;\ x^2 + x - 2 = 0$
   (c) $x^3 - x^2 - 14x + 24 = 0;\ x^2 + x - 6 = 0$
   (d) $x^3 + 2x^2 - 2x - 1 = 0;\ x^3 - 1 = 0$.

13.4 Discriminant. If the resultant of $f(x)$ and $f'(x)$ vanishes, then $f(x) = 0$ and $f'(x) = 0$ have a root in common. If this root is repeated m times in $f'(x) = 0$, it is repeated $m + 1$ times in $f(x) = 0$.

We now proceed to find this resultant:

If $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the roots of $f(x) = 0$, then

$$f(x) = a_0(x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \cdots (x - \alpha_n),$$

and

$$f'(x) = \frac{f(x)}{x - \alpha_1} + \frac{f(x)}{x - \alpha_2} + \cdots + \frac{f(x)}{x - \alpha_n};$$

whence

$$\begin{aligned}
f'(\alpha_1) &= a_0(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \cdots (\alpha_1 - \alpha_n) \\
f'(\alpha_2) &= a_0(\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3) \cdots (\alpha_2 - \alpha_n) \\
f'(\alpha_3) &= a_0(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)(\alpha_3 - \alpha_4) \cdots (\alpha_3 - \alpha_n) \\
&\cdots\cdots\cdots\cdots \\
f'(\alpha_n) &= a_0(\alpha_n - \alpha_1)(\alpha_n - \alpha_2) \cdots (\alpha_n - \alpha_{n-1}).
\end{aligned}$$

By §13.1, the resultant R of $f(x)$ and $f'(x)$ is

$$R = a_0^{n-1} f'(\alpha_1) f'(\alpha_2) \cdots f'(\alpha_n),$$

whence

$$\begin{aligned}
R &= a_0^{2n-1}(-1)^{1+2+\cdots+(n-1)}(\alpha_1 - \alpha_2)^2 \cdots (\alpha_{n-1} - \alpha_n)^2 \\
&= a_0^{2n-1}(-1)^{\frac{n(n-1)}{2}}(\alpha_1 - \alpha_2)^2 \cdots (\alpha_{n-1} - \alpha_n)^2.
\end{aligned}$$

The *discriminant* of an equation involving a single unknown is defined to be the "simplest" function of the coefficients, in a rational and integral form, whose vanishing expresses the condition for equal roots. Since, by definition, the discriminant of $f(x) = 0$ vanishes when two roots are equal, it follows that $\alpha_1 - \alpha_2$ must be a factor of the discriminant. But the discriminant is a fixed number for any given equation with constant coefficients. The given equation is unaltered, and, therefore, the discriminant must be unaltered, by an interchange of α_1 and α_2. Hence $(\alpha_1 - \alpha_2)^2$ must be a factor of the discriminant. The same reasoning applies to any two roots. It follows that

$$F = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2 \cdots (\alpha_1 - \alpha_n)^2(\alpha_2 - \alpha_3)^2 \cdots (\alpha_{n-1} - \alpha_n)^2$$

is a factor of the discriminant. The degree of F in any root is $2(n - 1)$. The function F is a symmetric function of the roots and so can be expressed rationally in terms of the coefficients of $f(x)$. When the substitutions $\sum \alpha_1 = -a_1/a_0, \cdots, \alpha_1\alpha_2 \cdots \alpha_n =$

$(-1)^n a_n/a_0$, are made, $2(n - 1)$ is the smallest power of a_0 which when multiplied by F will give a rational integral function of the coefficients. Hence the discriminant D may be identified by the expression

$$D = a_0^{2n-2}(\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2 \cdots (\alpha_1 - \alpha_n)^2(\alpha_2 - \alpha_3)^2 \cdots (\alpha_{n-1} - \alpha_n)^2.$$

By comparing the expressions for D and R where R is taken as the resultant of $f(x)$ and $f'(x)$, we find

$$R = (-1)^{\frac{n(n-1)}{2}} a_0 D$$

and

$$D = (-1)^{\frac{n(n-1)}{2}} \frac{1}{a_0} R.$$

It is seen that this definition of the discriminant is in harmony with the definition of the discriminant of a cubic given in chapter VII.

13.5 Discriminant of the quartic. In article 7.7, it was shown that if x_1, x_2, x_3, x_4 were the roots of the quartic

$$x^4 + 2px^3 + qx^2 + rx + s = 0, \tag{4}$$

and u_1, u_2, u_3 were the roots of the resolvent cubic

$$u^3 - qu^2 + (2pr - 4s)u + 4qs - 4p^2s - r^2 = 0, \tag{5}$$

then $u_1 = x_1x_2 + x_3x_4$; $u_2 = x_1x_3 + x_2x_4$; $u_3 = x_1x_4 + x_2x_3$. It now follows that

$$u_1 - u_2 = (x_1 - x_4)(x_2 - x_3), \quad u_1 - u_3 = (x_1 - x_3)(x_2 - x_4),$$

$u_2 - u_3 = (x_1 - x_2)(x_3 - x_4)$, and consequently

$$(u_1 - u_2)^2(u_1 - u_3)^2(u_2 - u_3)^2 = (x_1 - x_2)^2(x_1 - x_3)^2(x_1 - x_4)^2 (x_2 - x_3)^2(x_2 - x_4)^2(x_3 - x_4)^2.$$

But by definition the right-hand member of this last equation is the discriminant of the quartic (4), and the left-hand member is the discriminant of the resolvent cubic (5). Hence the discriminant of the quartic vanishes when the discriminant of the resolvent cubic vanishes. Furthermore, *the discriminant of the*

quartic (4) *can be computed by computing the discriminant of this particular resolvent cubic* (5).

In article 7.2, it was shown that the discriminant of

$x^3 + bx^2 + cx + d = 0$ is $18bcd - 4b^3d + b^2c^2 - 4c^3 - 27d^2$. (6)

Example 1: Compute the discriminant of $x^4 + x^2 + 1 = 0$:

The resolvent cubic is $u^3 - u^2 - 4u + 4 = 0$.

By (6), the discriminant of this cubic in u is 144; therefore, the discriminant of the original quartic is 144.

Example 2: Compute the discriminant of $x^4 - 2x^2 + 1 = 0$:

The resolvent cubic is $u^3 + 2u^2 - 4u - 8 = 0$.

By (6), the discriminant of the cubic is 0. Therefore the discriminant of the original quartic is 0. Hence the quartic must have a pair of equal roots: these roots are 1, 1, -1, -1.

Exercises

1. Compute the discriminant for the following equations:

(a) $x^3 + x^2 + x + 1 = 0$
(b) $x^3 + 4x^2 + 4x + 3 = 0$
(c) $x^3 + x + 1 = 0$
(d) $x^4 - 1 = 0$
(e) $x^4 + 1 = 0$
(f) $x^4 + 2x^3 - 3x^2 = 0$
(g) $x^3 - 5x + 4 = 0$
(h) $x^4 + 2x^3 - x + 5 = 0$

2. By computing the discriminant show that the following equations have equal roots:

(a) $x^3 + 2x^2 - 7x + 4 = 0$
(b) $x^3 + 7x^2 + 8x - 16 = 0$
(c) $x^3 + x^2 - 5x + 3 = 0$
(d) $x^4 - 3x^3 + x^2 + 4 = 0$
(e) $x^4 - x^3 - x + 1 = 0$
(f) $x^4 - 9x^2 + 4x + 12 = 0$
(g) $x^4 - 2x^3 - 27x^2 + 108 = 0$
(h) $x^4 - 12x^2 + 16x = 0$

3. Show that the discriminant for $ax^3 + bx^2 + cx + d = 0$ is $18abcd - 4b^3d + b^2c^2 - 4ac^3 - 27a^2d^2$.

4. Show that the discriminant for $y^3 + py + q = 0$ is $-4p^3 - 27q^2$.

5. Show that for $x^n + a_1x^{n-1} + \cdots + a_{n-2}x^2 = 0$ the discriminant vanishes and that, therefore, the equation has equal roots.

6. Show that for (a) $x^n - 1 = 0$; (b) $x^n + 1 = 0$, the discriminant does not vanish and that, therefore, the equation cannot have equal roots.

Chapter XIV

RULER AND COMPASS CONSTRUCTIONS

14.1 Introduction. The ancient Greeks proposed a number of problems in construction in which only a ruler and a compass were to be used. Among these were the problem of:

1. The duplication of a cube.
2. The trisection of any angle.
3. The construction of a regular polygon of any number of sides.

The ruler must be a rigid body with a straightedge and without division marks. Given two points we can with the straightedge draw the line segment connecting the points and prolong a line segment indefinitely. With the compass one can draw a circle whose center is any given point and radius any given line segment.* That is, the fundamental constructions are

1. Drawing the line segment between two points.
2. Extending a line segment.
3. Drawing a circle with a given point as center and a given line segment as radius.

All other constructions consist of two or more applications of these fundamental constructions.

New points, lines, and line segments may be found as

1. The intersection of two lines.
2. The intersection of a line and a circle.
3. The intersection of two circles.

In all that follows it is assumed that the unit segment is given.

14.2 Graphical solution of a quadratic equation.
If the quadratic equation

$$x^2 - ax + b = 0 \tag{1}$$

* Initially one need suppose only that the compass serves to draw a circle with given center and passing through a given point. Euclid showed how one could then construct the circle with given center and with radius equal to a given line segment.

has real roots, and the coefficients a and b are real and can be constructed, then these real roots can be constructed with ruler and compass as follows:

Draw a circle whose diameter is the line joining the points $A(0, 1)$ and $B(a, b)$. The center C of this circle has the coordinates $\left(\frac{a}{2}, \frac{b+1}{2}\right)$. The radius $CB = \sqrt{\frac{a^2}{4} + \frac{(b-1)^2}{4}}$. Hence the equation of the circle is

$$\left(x - \frac{a}{2}\right)^2 + \left(y - \frac{b+1}{2}\right)^2 = \frac{a^2 + (b-1)^2}{4}. \tag{2}$$

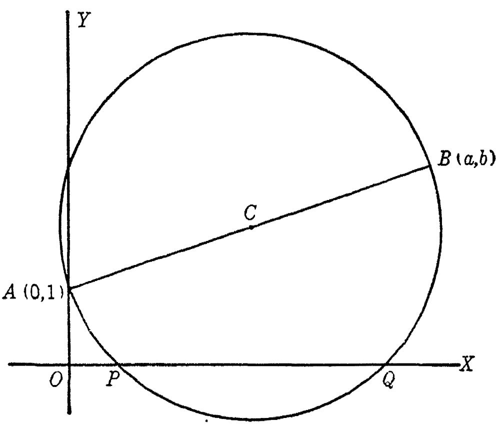


Fig. 33

When $y = 0$, this equation reduces to (1), which proves the statement that the roots, when real, can be constructed by ruler and compass. The roots of (1) are represented by OP and OQ.

Exercises

Solve graphically:

1. $x^2 - 3x + 2 = 0$
2. $x^2 - 6x + 5 = 0$
3. $x^2 - 4x + 4 = 0$
4. $x^2 + 4x + 4 = 0$
5. $x^2 - 5x - 6 = 0$
6. $x^2 + 6x - 7 = 0$

14.3 Elementary constructions with ruler and compass. Given two line segments AB and CD, fig. 34. With B as a center and

a radius CD draw a circle. Extend AB to meet the circle at P. Then

$$AP = AB + CD$$

and
$$AQ = AB - CD.$$

Given two line segments a and b, fig. 35, draw any two intersecting lines as AC and AD. Lay off on AC the segment $AB = 1$ and $BC = a$. On AD lay off $AE = b$. This can be done with

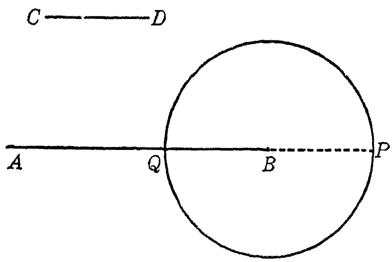


Fig. 34

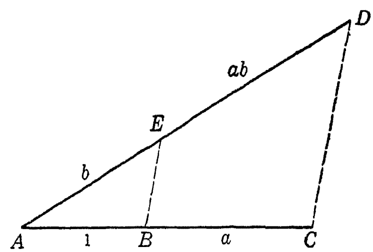


Fig. 35

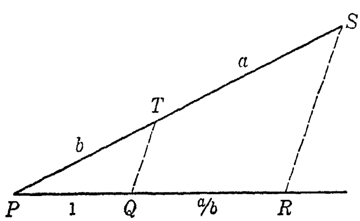


Fig. 36

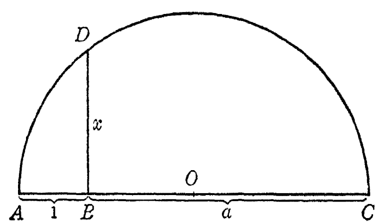


Fig. 37

the compasses. Join B and E. Our straightedge and compass are sufficient to construct CD parallel to BE. Then $ED = ab$.

Given two line segments a and b we find their sum PS, fig. 36. Draw any other line through P, as PR. Lay off with the compass $PQ = 1$. Draw RS parallel to QT and intersecting PT in S. Then $QR = a/b$.

On any line lay off $AB = 1$, and $BC = a$. Find the middle point O of AC. With O as center and radius OC draw a semicircle. Erect BD perpendicular to AC at B. Then $x = \sqrt{a}$. That is, this construction solves the equation $x^2 - a = 0$. (*See* fig. 37.)

Hence, if the finite set of numbers $1, a, b, \cdots$ correspond to given line segments and are combined a finite number of times by the operations $+, -, \times, \div, \sqrt{\ }$, to produce a real number n, the corresponding line segments can be combined geometrically by ruler and compass to produce the line segment corresponding to the number n.

14.4 Criterion for constructibility. *First.* If at each step in the finite process there are no irrational operations other than extracting a real square root, then the constructions can be made; since we can perform all constructions involving the rational operations $+, -, \times, \div$, and we have seen how we can extract the square root of a *real positive number*, which is itself constructible.

Second. Suppose the construction is possible. Then the straight lines and circles drawn in making the construction are obtained either from points and line segments initially given or obtained from them as

1. the intersection of two lines
2. the intersection of a line and a circle
3. the intersection of two circles.

In order to formulate algebraically every construction by means of a ruler and compass we must show how to find algebraically the intersections just mentioned.

The equations of any two nonparallel lines may be taken as

$$A_1x + B_1y = C_1, \qquad A_2x + B_2y = C_2, \qquad A_1B_2 - A_2B_1 \neq 0.$$

The coordinates of their point of intersection are

$$x = \frac{C_1B_2 - C_2B_1}{A_1B_2 - A_2B_1}, \qquad y = \frac{A_1C_2 - A_2C_1}{A_1B_2 - A_2B_1} \qquad (A_1B_2 - A_2B_1 \neq 0),$$

which are rational functions of the coefficients *and hence constructible* if the given lines are regarded as themselves constructed.

Let a constructible straight line $Ax + By = C$ intersect the constructible circle

$$(x - c)^2 + (y - d)^2 = r^2$$

with the center (c, d) and radius r. To find the coordinates of the points of intersection, eliminate y between the two equations, and obtain a quadratic for the determination of x. The coefficients of this quadratic are rational functions of the parameters

and hence can be constructed. The quadratic has real coefficients (themselves constructible) and real roots and hence the roots can be constructed, and also y. Thus x (and also y) involve no irrationality other than a real square root, besides the irrationalities present in A, B, C, c, d, r.

The intersection of two nonconcentric circles

$$\begin{aligned} x^2 + y^2 + A_1x + B_1y + C_1 &= 0 \quad (A_1 - A_2 \text{ and } B_1 - B_2 \\ x^2 + y^2 + A_2x + B_2y + C_2 &= 0 \qquad \text{not both zero)} \end{aligned}$$

is obtained from the intersections of either of the circles with their common chord

$$(A_1 - A_2)x + (B_1 - B_2)y + C_1 - C_2 = 0,$$

and hence reduces to the preceding case.

We have now proved the following theorem:

Theorem: A necessary and sufficient condition that a line segment of length x be constructible by ruler and compass from given line segments of lengths $a_1, \cdots, a_n$ is that the number x can be derived from the given numbers $a_1, \cdots, a_n$ by a finite number of the rational operations $+$, $-$, $\times$, $\div$ and extractions of real square roots.

14.5 Simplification of radicals. We agree to make all possible simplifications in any such expression x containing radicals. Such simplifications are illustrated by the following examples:

We agree to write an expression containing radicals in such a way that no one of its radicals can be expressed as a rational function in terms of the other radicals, with coefficients which are either themselves rational or if irrational are of lower order of irrationality in the sense to be explained below. Thus $\sqrt{3} + \sqrt{5} + \sqrt{15}$ would be replaced by $\sqrt{3} + \sqrt{5} + \sqrt{3} \cdot \sqrt{5}$.

We agree to write $\sqrt{3 + \sqrt{13 + 4\sqrt{3}}}$ in its simpler form $\sqrt{3} + 1$.

14.6 Order of a radical: A radical expression of the form $\sqrt{5 + 3\sqrt{2}}$ is said to be a radical of order 2. One radical sign is placed over an expression containing a radical.

A radical is said to be of order n if a radical sign is placed over an expression containing a radical of order $n - 1$, and if this radical of order n cannot be expressed as a rational function, with

rational coefficients, of a finite number of radicals of order less than n.

We shall suppose that in a number x containing t terms of order n, these reductions have been made so that no one of the radicals of order r ($r \leq n$) can be expressed as a rational function, with rational coefficients, of the remaining radicals of order r and the radicals of order lower than r.

14.7 Normal form of x. Suppose $\sqrt{N}$ one of the terms of x of order n. Then at worst

$$x = \frac{A + B\sqrt{N}}{C + D\sqrt{N}} \qquad (C, D \neq 0),$$

where A, B, C, D may contain radicals of order n. Multiply both numerator and denominator of the fraction by $C - D\sqrt{N}$. We have then, on simplifying, the relation of the form

$$x = a + b\sqrt{N}$$

where a and b contain no more than $t - 1$ terms of order n. We may then put

$$a = a_1 + a_2\sqrt{N_1}, \qquad b = b_1 + b_2\sqrt{N_1}$$

where $\sqrt{N_1}$ is a radical of order n at most. Then

$$x = (a_1 + a_2\sqrt{N_1}) + (b_1 + b_2\sqrt{N_1})\sqrt{N}.$$

Proceed in this way until all of the radicals of order n appear explicitly. Proceed in a similar way with the different terms of order $n - 1$ which occur in the coefficients and in N, N_1, etc. Then pass on to terms of lower order. We finally obtain x in the form of a rational integral linear function of each individual radical all of which occur explicitly. x is then said to be reduced to the normal form.

14.8 Theorem. *If x, the number to be constructed, depends only upon a finite number of rational operations and extraction of square roots of real positive numbers, it is a root of an equation with rational coefficients, whose degree is a power of* 2.

Let m be the number of independent radicals in x. By changing the sign of one or more of these radicals we obtain 2^m different values $x_1 = x, x_2, x_3, \cdots, x_{2^m}$, called mutually *conjugate values.*

These values may not all be distinct. For example, if $x_1 = \sqrt{3 + \sqrt{2}} + \sqrt{3 - \sqrt{2}}$, of the 8 conjugate values only 4 are distinct, namely:

$$x_1; \qquad x_2 = \sqrt{3 + \sqrt{2}} - \sqrt{3 - \sqrt{2}};$$
$$x_3 = -\sqrt{3 + \sqrt{2}} + \sqrt{3 - \sqrt{2}}; x_4 = -\sqrt{3 + \sqrt{2}} - \sqrt{3 - \sqrt{2}}.$$

Form the equation

$$P(x) = (x - x_1)(x - x_2) \cdots (x - x_{2^m}).$$

Clearly *this equation is of degree* 2^m. It may have multiple roots. Certainly if all of the 8 conjugates of $x_1 = \sqrt{3 + \sqrt{2}} + \sqrt{3 - \sqrt{2}}$ are used, each root would be a double root.

The coefficients of $P(x)$ *are rational.* For, if we change the sign of any radical, this causes two of the factors in $P(x)$ to be interchanged and so $P(x)$ remains unchanged. This means that the coefficients in $P(x)$ remain unchanged. This can happen only if each radical occurs in the coefficients in its squared form only. Hence the coefficients of $P(x)$ are rational.

Illustration 1: Let $x = \sqrt{2 + \sqrt{5}}$. There are $2^2 = 4$ conjugate values. Then x should satisfy an equation of degree 4 with rational coefficients.

Square

$$x^2 = 2 + \sqrt{5} \quad \text{or} \quad x^2 - 2 = \sqrt{5}.$$

Square again

$$x^4 - 4x^2 + 4 = 5$$

or

$$x^4 - 4x^2 - 1 = 0.$$

Illustration 2: Let $x = \sqrt{2} + \sqrt{5}$. There are $2^2 = 4$ conjugate values: $\sqrt{2} + \sqrt{5}$; $\sqrt{2} - \sqrt{5}$; $-\sqrt{2} + \sqrt{5}$; $-\sqrt{2} - \sqrt{5}$. Then x should satisfy an equation of degree 4 with rational coefficients.

$$x = \sqrt{2} + \sqrt{5}$$

Squaring

$$x^2 = 7 + 2\sqrt{10} \quad \text{or} \quad x^2 - 7 = 2\sqrt{10};$$

square again

$$x^4 - 14x^2 + 49 = 40$$

or

$$x^4 - 14x^2 + 9 = 0.$$

Exercises

Form the equation $P(x) = 0$ which has the following numbers as roots:

1. $x = \sqrt{2 + \sqrt{3}}$
2. $x = \sqrt{2} + \sqrt{3}$
3. $x = \sqrt{2 + \sqrt{3}} + \sqrt{2 - \sqrt{3}}$
4. $x = \sqrt{5 + \sqrt{3}}$
5. $x = \sqrt{3} + \sqrt{5}$
6. $x = \sqrt{2 + \sqrt{5}} + \sqrt{2 - \sqrt{5}}$

14.9 Theorem. *If any one of the conjugates $x_1, \cdots, x_{2^m}$ satisfies an equation $f(x) = 0$, with rational coefficients, then all of the conjugates satisfy that equation.*

Let $x_1 = a + b\sqrt{N}$ where $\sqrt{N}$ is a radical of maximum order n, while a and b do not involve $\sqrt{N}$ but may depend upon other radicals of order n and radicals of lower order. Then $x_2 = a - b\sqrt{N}$ will be one of the conjugates of x_1.

Now $f(x_1)$ may be given the form

$$f(x_1) = A_1 + B_1\sqrt{N}$$

where A_1 and B_1 do not involve $\sqrt{N}$. Here $f(x_1)$ can vanish only if $A_1 = B_1 = 0$. Otherwise if $A_1 \neq 0$ and $B_1 \neq 0$ we would have $\sqrt{N} = -A_1/B_1$. That is, $\sqrt{N}$ could be expressed rationally as a function of terms of order n, and of terms of lower order, contrary to the hypotheses of the independence of the radicals. If $A_1 \neq 0$ while $B_1 = 0$, or $A_1 = 0$ while $B_1 \neq 0$, the relation $A_1 + B_1\sqrt{N} = 0$ could not hold. It follows that

$$f(x_2) = A_1 - B_1\sqrt{N} = 0.$$

Hence, the conjugate x_2 is a root of $f(x) = 0$. We have now proved that if x_1 satisfies $f(x) = 0$, then all of the conjugates derived from x_1, by changing the signs of one or more radicals of order n, also satisfy $f(x) = 0$.

The proof that *any* conjugate x_i is a root of $f(x) = 0$ follows along similar lines. For simplicity let us assume that x_1 depends on only two radicals of order n, namely $\sqrt{N}$ and $\sqrt{N_1}$. Then

$$f(x_1) = A_2 + B_2\sqrt{N} + C_2\sqrt{N_1} + D_2\sqrt{N}\sqrt{N_1} = 0. \quad (3)$$

If x_1 depended on more than two radicals of order n, we should only have to add to the above expression for $f(x_1)$ more terms of like nature.

Equation (3) is possible only when $A_2 = B_2 = C_2 = D_2 = 0$. Otherwise $\sqrt{N}$ could be expressed rationally as a function of radicals of order n and of radicals of lower order, which is contrary to the hypothesis.

Let $\sqrt{N_{11}}$, $\sqrt{N_{12}}$, $\cdots$ be the radicals of order $n - 1$ in x_1. These radicals occur in A_2, B_2, C_2, D_2. If for simplicity we consider only two radicals $\sqrt{N_{11}}$, $\sqrt{N_{12}}$, we have

$$A_2 = A_3 + B_3\sqrt{N_{11}} + C_3\sqrt{N_{12}} + D_3\sqrt{N_{11}}\sqrt{N_{12}} \qquad (4)$$

and analogous expressions for B_2, C_2, D_2. On account of the independence of the radicals, we must have

$$A_3 = B_3 = C_3 = D_3 = 0.$$

Equations (4), and hence also $f(x) = 0$, are satisfied when x_1 is replaced by any one of its conjugates obtained by changing the signs of $\sqrt{N_{11}}$ and $\sqrt{N_{12}}$. Thus we see that $f(x) = 0$ is satisfied when x_1 is replaced by any one of its conjugates obtained by changing the signs of the radicals of order $n - 1$. The same reasoning applies to the radicals of order $n - 2$, $n - 3$, $\cdots$. Hence the theorem is proved.

Illustration. Let $x = -\sqrt{2} + \sqrt{1 + \sqrt{3}}$. There are $2^3 = 8$ conjugate values. $a = -\sqrt{2}$ is a radical of order 1. $b = 1$. $\sqrt{N} = \sqrt{1 + \sqrt{3}}$ is a radical of order 2. We have

$$x + \sqrt{2} = \sqrt{1 + \sqrt{3}}.$$

Squaring

$$x^2 + 2\sqrt{2}x + 2 = 1 + \sqrt{3}$$

or

$$x^2 + 2\sqrt{2}x + 1 = \sqrt{3}.$$

Square and collect terms. We have

$$x^4 + 10x^2 - 2 = -4\sqrt{2}(x^3 + x).$$

Square again and collect terms. We have

$$x^8 - 12x^6 + 32x^4 - 72x^2 + 4 = 0.$$

Exercises

Form the equations with rational coefficients which have the following roots:

1. $x = -\sqrt{3} + \sqrt{1 + \sqrt{2}}$.

Ans. $x^8 - 16x^6 + 68x^4 - 128x^2 + 4 = 0$.

2. $x = -\sqrt{5} + \sqrt{4 + \sqrt{3}}$.

Ans. $x^8 - 36x^6 + 320x^4 - 168x^2 + 4 = 0$.

3. $x = -1 - \sqrt{2} + \sqrt{3 + \sqrt{3}}$.
Ans. $x^8 + 8x^7 + 8x^6 - 64x^5 - 134x^4 + 40x^3 + 248x^2 + 160x + 25 = 0$.

14.10 The equation $\phi(x) = 0$ of lowest degree satisfied by x_1 and its distinct conjugates.

I. $\phi(x)$ *is irreducible.* Suppose $\phi(x)$ were reducible, then by definition

$$\phi(x) = \alpha(x)\beta(x),$$

where $\alpha(x)$ and $\beta(x)$ have as their coefficients rational numbers.* Since $\phi(x_1) = 0$, either $\alpha(x_1) = 0$ or $\beta(x_1) = 0$ or both. But if $\alpha(x) = 0$ [$\beta(x) = 0$] admits x_1 as a root, it admits all of its conjugates as roots. Then we would have an equation $\alpha(x) = 0$ [$\beta(x) = 0$], of lower degree than $\phi(x)$, satisfied by x_1 and its distinct conjugates. This would be contrary to our assumption that $\phi(x) = 0$ was the equation of lowest degree satisfied by x_1 and its distinct conjugates. Hence $\phi(x)$ is irreducible.

II. $\phi(x) = 0$ *has no multiple roots.* For if it had multiple roots, $\phi(x)$ and $\phi'(x)$ would have as highest common factor a polynomial with rational coefficients and of positive degree, and hence $\phi(x)$ would be reducible.

III. $\phi(x) = 0$ has no roots other than x_1 and its distinct conjugates. Otherwise $P(x)$ and $\phi(x)$ would have as highest common divisor a polynomial of positive degree and $\phi(x)$ would not then be irreducible.

IV. Since $\phi(x) = 0$ has no multiple roots and $\phi(x_1) = 0$, then the only roots of $\phi(x) = 0$ are x_1 and its distinct conjugates.

V. $\phi(x) = 0$ is (save for a trivial constant multiplier) the only irreducible equation with rational coefficients satisfied by x_1 and its distinct conjugates. If there were another $f(x) = 0$, then $f(x)$ would have $\phi(x)$ as a factor and accordingly would be either a constant multiple of $\phi(x)$ or would be reducible.

VI. There cannot be two different irreducible equations of the same degree. Otherwise, by eliminating the term of highest degree

* If any polynomial $\phi(x)$ with rational coefficients is the product of two polynomials $\alpha(x)$ and $\beta(x)$ both of positive degree, and if one of these has rational coefficients, the other will, also. This may be seen from the nature of the process of division by which the coefficients of the second factor polynomial may be identified.

we could obtain an equation of lower degree satisfied by x_1 and its distinct conjugates. Then $\phi(x) = 0$ would not be the equation of lowest degree satisfied by the x_i's.

Hence $\phi(x) = 0$ is *the* irreducible equation of lowest degree satisfied by x_1 and its distinct conjugates.

14.11 $P(x)$ is an integral power of $\phi(x)$. $P(x) = 0$ and $\phi(x) = 0$ have x_1 and its distinct conjugates as their only roots, and $\phi(x) = 0$ has no multiple roots. Then

$$P(x) = P_1(x) \cdot \phi(x)$$

and $P_1(x)$ has rational coefficients. If $P_1(x)$ is not constant, then its roots are also roots of $P(x)$. That is, x_1 or one of its conjugates is a root of $P_1(x) = 0$. But in that case all of the x_i's are roots of $P_1(x) = 0$. Then

$$P_1(x) = P_2(x) \cdot \phi(x).$$

Again, reasoning in the same way, either $P_2(x)$ is a constant or of the form $P_2(x) = P_3(x) \cdot \phi(x)$. At each step the degree of the quotient is decreased by one. So after a finite number of steps we reach an equation of the form

$$P_{s-1}(x) = C\phi(x), \qquad [C, \text{ a constant } \neq 0].$$

Hence

$$P(x) = C[\phi(x)]^s.$$

If the degree of $\phi(x)$ *is* r, then

$$2^n = r \cdot s.$$

Therefore r is also a power of 2.

Hence we have the following theorem:

Theorem: *The degree of the irreducible equation satisfied by a number* **x** *obtained by a finite number of rational operations and extraction of square roots is a power of 2.*

Since there is only one irreducible equation satisfied by all of the x_i's, we have the following theorem:

Theorem: *If the degree of an irreducible equation is not a power of two, the equation cannot be solved by a finite number of rational operations and extraction of square roots.*

Illustration: Let $x = \sqrt{3+\sqrt{2}} + \sqrt{3-\sqrt{2}}$. There are $2^3 = 8$ conjugates. The equation of degree 8 that has all of these 8 conjugates as roots is

$$x^8 - 24x^6 + 16x^4 - 48x^2 + 64 = (x^4 - 12x^2 + 8)^2 = 0.$$

In this case only 4 of the conjugates are distinct. The irreducible equation satisfied by all of the conjugates is $\phi(x) = x^4 - 12x^2 + 8 = 0$.

Exercises

Form the irreducible equation $\phi(x) = 0$, with rational coefficients, which is satisfied by the following value of x:

1. $x = \sqrt{5+\sqrt{7}}$
2. $x = \sqrt{2} + \sqrt{5}$
3. $x = \sqrt{3+\sqrt{5}} + \sqrt{3-\sqrt{5}}$
4. $x = \sqrt{2} + \sqrt{3} + \sqrt{5}$
5. $x = \sqrt{1+\sqrt{2}} + \sqrt{2+\sqrt{2}}$
6. $x = \sqrt{1+2\sqrt{2}} + \sqrt{1+3\sqrt{2}}$
7. $x = \sqrt{3+\sqrt{3+\sqrt{2}}}$
8. $x = \sqrt{5} + \sqrt{3+\sqrt{2}}$
9. $x = \sqrt{-2+\sqrt{3+\sqrt{1+\sqrt{2}}}}$
10. $x = -2 - \sqrt{2} + \sqrt{3+\sqrt{3}}$
11. $x = -1 - \sqrt{3} - \sqrt{1+\sqrt{2}}$
12. $x = 1 + \sqrt{2} - \sqrt{3}$
13. $x = 1 + \sqrt{2} - \sqrt{5}$

14.12 Duplication of the cube: The problem is to find the edge x of a cube whose volume is twice that of a given cube. Analytically this is equivalent to solving the equation $x^3 = 2$.

If this equation was reducible, it would have at least one rational linear factor. Any rational root is an integral divisor of 2. By trial ± 1, ± 2 are not roots. Hence $x^3 - 2 = 0$ is irreducible. Since $x^3 - 2 = 0$ is an irreducible equation whose degree is not a power of 2, it does not have a constructible root. Hence the desired construction is impossible with ruler and compass.

14.13 Trisection of an angle. To show that not all angles can be trisected, it is sufficient to show that $120°$ cannot be trisected.

An angle of 120° can be trisected if and only if we can construct sin 40°. For, given an angle of 40° we can construct a right triangle, with unit hypotenuse, that contains the angle. The length of the side opposite the angle of 40° is numerically equal to sin 40°. If we have given sin 40°, at one end A of the line segment whose length is sin 40° construct a perpendicular. With the other end B as center of a circle, with unit radius, construct an arc intersecting the perpendicular in C. Then $\angle ACB = 40°$.

From trigonometry we have $\cos 120° = 4 \cos^3 40° - 3 \cos 40°$. Multiply each term by 2. Replace $2 \cos 40°$ by x and $2 \cos 120°$ by -1; we have

$$x^3 - 3x + 1 = 0. \tag{5}$$

If this equation was reducible,* it would have a rational linear factor. Any rational root is an integer and a divisor of 1. By trial ± 1 are not roots. Hence (5) is irreducible. Since (5) is an irreducible equation whose degree is not a power of 2, it does not have a constructible root. So it is impossible to trisect some angles with ruler and compass, and no Euclidean construction for trisecting all angles is possible.

14.14 Regular polygon of 9 sides. In a regular polygon of 9 sides, one of the equal sides subtends an angle of 40° at the center. By the previous article we have seen that $x = 2 \cos 40°$ cannot be constructed by ruler and compass. Then $x/2 = \cos 40°$ cannot be so constructed. Hence 40° cannot be so constructed. Thus we see that it is not possible to construct with ruler and compass a regular polygon of 9 sides.

14.15 Regular polygon of 7 sides. In a regular polygon of 7 sides, one of the equal sides subtends, at the center, an angle A of 360/7 degrees. If we could construct this angle we could construct a line segment of length $x = 2 \cos A$.

Since

$$7A = 360°, \qquad \cos 4A = \cos 3A.$$

But

$$2 \cos 4A = 2(2 \cos^2 2A - 1) = 4(2 \cos^2 A - 1)^2 - 2 = (x^2 - 2)^2 - 2$$

$$2 \cos 3A = 2(4 \cos^3 A - 3 \cos A) = x^3 - 3x.$$

* The reader should not confuse this use of "reducible" with that in the "irreducible case," nor with that in the "reduced cubic."

Therefore

$$(x^2 - 2)^2 - 2 = x^3 - 3x.$$

Whence

$$x^4 - 4x^2 + 2 - x^3 + 3x = (x - 2)(x^3 + x^2 - 2x - 1) = 0.$$

If $x = 2$, then would $\cos A = 1$, whereas $\cos \frac{360°}{7} < 1$. Hence

$$x^3 + x^2 - 2x - 1 = 0. \tag{6}$$

If this equation were reducible, it would have a rational linear factor. Any rational root would be an integer. By trial ± 1 are not roots. Hence (6) is irreducible. Since (6) is an irreducible equation whose degree is not a power of 2, it does not have a constructible root. Hence, with ruler and compass, it is impossible to construct a regular polygon of 7 sides.

Exercises

1. Prove that the angles 10°, 20°, 50°, 70°, 80° cannot be constructed with ruler and compass alone.

(In some cases use $\sin 3A = 3 \sin A - 4 \sin^3 A$.)

2. Prove that every real root of $x^4 + px^2 + q = 0$ can be constructed with ruler and compass, given lines whose lengths are p and q.

3. Prove that it is possible to construct with ruler and compass alone the legs of a right triangle, given the length of the hypotenuse and the area.

14.16 Regular polygons of 5 and 10 sides.

$$R = \cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}$$

is a root of $y^5 = 1$ and hence of

$$y^4 + y^3 + y^2 + y + 1 = 0. \tag{7}$$

By the substitution $x = y + \dfrac{1}{y}$, (7) becomes

$$x^2 + x - 1 = 0,$$

which has the root $R + \dfrac{1}{R} = 2 \cos \dfrac{2\pi}{5} = \frac{1}{2}(\sqrt{5} - 1)$.

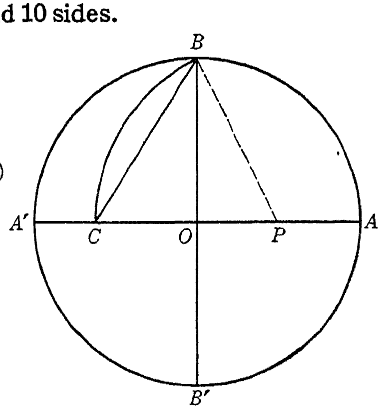


Fig. 38

In a circle of radius unity and center O, draw two perpendicular diameters AA' and BB'. Find the midpoint P of OA. With P as a center and radius PB draw a circle cutting AA' at C. Then BC and OC are the sides of the inscribed regular polygons of 5 and 10 sides respectively.

We have

$$PB = \sqrt{5}/2;$$

$$OC = \tfrac{1}{2}(\sqrt{5} - 1) = 2 \cos \frac{2\pi}{5} = 2 \sin 18°;$$

$$BC = \sqrt{1 + \overline{OC}^2} = \tfrac{1}{2}\sqrt{10 - 2\sqrt{5}}.$$

Whence $\overline{BC}^2 = \tfrac{1}{4}(10 - 2\sqrt{5}) = 2\left(1 - \cos \frac{2\pi}{5}\right) = (2 \sin 36°)^2.$

14.17 Regular polygon of 17 sides. The possibility of constructing a regular polygon of 17 sides is proved if we show that $\cos \frac{2\pi}{17}$ can be constructed. Consider the equation

$$x^{16} + x^{15} + x^{14} + \cdots + x^2 + x + 1 = 0.$$

One root is $R = \cos \frac{2\pi}{17} + i \sin \frac{2\pi}{17}$.

Gauss arranged the roots in the order

$$R,\ R^3,\ R^9,\ R^{10},\ R^{13},\ R^5,\ R^{15},\ R^{11},\ R^{16},\ R^{14},\ R^8,\ R^7,\ R^4,\ R^{12},\ R^2,\ R^6,$$

each of which is the cube of the preceding, the first being the cube of the last.

Set

$$y_1 = R + R^9 + R^{13} + R^{15} + R^{16} + R^8 + R^4 + R^2$$

$$y_2 = R^3 + R^{10} + R^5 + R^{11} + R^{14} + R^7 + R^{12} + R^6.$$

Then $y_1 + y_2 = -1$, $y_1 y_2 = -4$, and each y satisfies the equation $y^2 + y - 4 = 0$ whose roots are

$$y = \frac{\pm\sqrt{17} - 1}{2}. \tag{8}$$

But

$$y_1 = (R + R^{16}) + (R^2 + R^{15}) + (R^4 + R^{13}) + (R^8 + R^9)$$

$$= 2\cos\frac{2\pi}{17} + 2\cos\frac{4\pi}{17} + 2\cos\frac{8\pi}{17} + 2\cos\frac{16\pi}{17} > 0$$

$$\therefore y_1 = \frac{\sqrt{17}-1}{2}, \qquad y_2 = \frac{-\sqrt{17}-1}{2}.$$

Now set

$$z_1 = R + R^{13} + R^{16} + R^4 = 2\cos\frac{2\pi}{17} + 2\cos\frac{8\pi}{17} > 0$$

$$z_2 = R^9 + R^{15} + R^8 + R^2 = 2\cos\frac{4\pi}{17} + 2\cos\frac{16\pi}{17} < 0.$$

Then $z_1 + z_2 = y_1$, $z_1 z_2 = -1$, and each z satisfies the equation.

$$z^2 - y_1 z - 1 = 0, \text{ whose roots are} \tag{9}$$

$$z_1 = \frac{\sqrt{17}-1}{4} + \frac{\sqrt{34 - 2\sqrt{17}}}{4};$$

$$z_2 = \frac{\sqrt{17}-1}{4} - \frac{\sqrt{34 - 2\sqrt{17}}}{4}.$$

Now set

$$w_1 = R^3 + R^5 + R^{14} + R^{12} = 2\cos\frac{6\pi}{17} + 2\cos\frac{10\pi}{17} > 0$$

$$w_2 = R^{10} + R^{11} + R^7 + R^6 = 2\cos\frac{12\pi}{17} + 2\cos\frac{14\pi}{17} < 0.$$

Then $w_1 + w_2 = y_2$, $w_1 w_2 = -1$, and each w satisfies the equation

$$w^2 - y_2 w - 1 = 0, \text{ whose roots are}$$

$$w_1 = \frac{-\sqrt{17}-1}{4} + \frac{\sqrt{34 + 2\sqrt{17}}}{4},$$

$$w_2 = \frac{-\sqrt{17}-1}{4} - \frac{\sqrt{34 + 2\sqrt{17}}}{4}.$$

Now set

$$u_1 = R + R^{16} = 2\cos\frac{2\pi}{17}, \qquad u_2 = R^4 + R^{13} = 2\cos\frac{8\pi}{17}.$$

Then

$$u_1 > u_2 > 0, \qquad u_1 + u_2 = z_1, \qquad u_1 u_2 = w_1.$$

Whence each u satisfies the equation

$$u^2 - z_1 u + w_1 = 0 \text{ whose roots are} \tag{11}$$

$$u_1 = \frac{z_1 + \sqrt{z_1^2 - 4w_1}}{2} = 2\cos\frac{2\pi}{17}, \qquad u_2 = \frac{z_1 - \sqrt{z_1^2 - 4w_1}}{2},$$

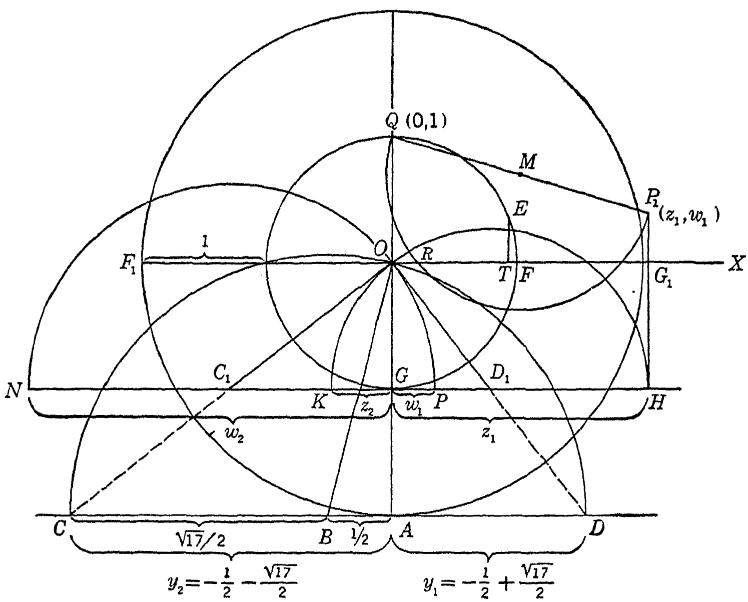


Fig. 39

and we see that $u_1/2 = \cos\frac{2\pi}{17}$ can be obtained by a finite number of rational operations and extraction of square roots of real numbers. Hence the regular polygon of 17 sides is constructible.

Construction

With O as center draw a circle with unit radius and another with radius two. Construct two perpendicular diameters OA and OF_1. Draw CD tangent to the larger circle at A and C_1D_1 to the smaller circle at G on OA. Take $AB = -\frac{1}{2}$. Then $OB = \sqrt{17}/2$.

With B as a center and BO as a radius construct a circle cutting AB at C and D. Connect C and D with O through C_1 and D_1.

$$AC = y_2 = -\frac{1}{2} - \frac{\sqrt{17}}{2} \quad \text{and} \quad AD = y_1 = -\frac{1}{2} + \frac{\sqrt{17}}{2}.$$

$$GD_1 = \tfrac{1}{2}AD, \qquad GC_1 = \tfrac{1}{2}AC.$$

$$OD_1 = \sqrt{\overline{OG}^2 + \overline{GD}_1^2} = \sqrt{1 + \tfrac{1}{4}y_1^2};$$

$$OC_1 = \sqrt{\overline{OG}^2 + \overline{GC}_1^2} = \sqrt{1 + \tfrac{1}{4}y_2^2}.$$

With D_1 as a center and OD_1 as a radius draw a circle cutting GD_1 at H and K. With C_1 as a center and OC_1 as a radius draw a circle cutting GD_1 at N and P. Then $GN = w_2$, $GP = w_1$, $GK = z_2$, $GH = z_1$, when taken as in fig. 39. In detail

$$GH = GD_1 + D_1H = GD_1 + OD_1 = \tfrac{1}{2}y_1 + \sqrt{1 + \tfrac{1}{4}y_1^2} = z_1,$$

$$GP = C_1P - C_1G = C_1O - C_1G = \tfrac{1}{2}y_2 + \sqrt{1 + \tfrac{1}{4}y_2^2} = w_1,$$

We now proceed to construct the roots of equation (11).

Erect a perpendicular to GH at H. Designate by G_1 the intersection of this perpendicular with OX. Lay off on HG_1 produced a segment $G_1P_1 = GP$. Then if O is the origin of coordinates and OX is the x-axis, the coordinates of P_1 are (z_1, w_1). The coordinates of Q are $(0, 1)$. Let M be the middle point of the line joining Q and P_1. Draw the semicircle with M as center and QM as radius. This will intersect OX in two points R and S. By §14.2 OR and OS are the roots of (11). The larger root

$$OS = u_1 = 2 \cos \frac{2\pi}{17}.$$

At the midpoint T of OS erect a perpendicular cutting the unit circle at E. Then $OT = \cos \frac{2\pi}{17}$ and $\angle TOE = \frac{2\pi}{17}$. Hence the chord EF is one side of the regular polygon of 17 sides inscribed in the unit circle, where F is the point $(1, 0)$.

Exercises

1. Show that one can construct regular polygons of 3, 4, 6, 8, 12, 15, 16, 20, 24, 34 sides.

2. Show that with ruler and compass alone one cannot trisect an angle whose sine or cosine is 1/2, 1/3, 2/3, 1/4, 3/4.

3. Show that with ruler and compass alone one cannot construct a regular polygon of 18 sides; $9 \cdot 2^n$ sides (n an integer).

4. The accompanying figure* shows an elementary construction for a regular inscribed pentagon and decagon. Only the Pythagorean theorem is necessary in the demonstration.

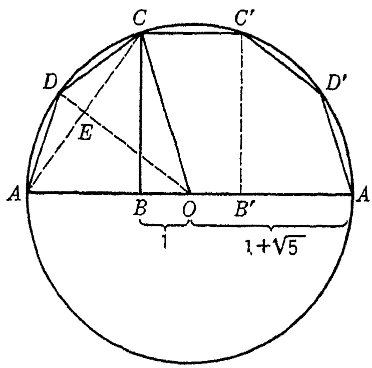


Fig. 40

Verify that if

$$OC = 1 + \sqrt{5}, \qquad OB = 1, \quad \text{we have}$$

$$AE = \tfrac{1}{2}\sqrt{10 + 2\sqrt{5}}; \qquad OE = \tfrac{1}{2}(3 + \sqrt{5});$$

$$AD = DC = CC' = C'D' = D'A' = 2.$$

Given $BC \perp OA$; $OD \perp AC$; $OB' = 1$; $B'C' \perp OA'$; D' bisects arc $A'C'$.

References

Klein: *Vorträge über ausgewählte Fragen der Elementargeometrie*, Leipzig, 1895. English translation by Beman and Smith, 1897. Second edition, revised and enlarged with notes by R. C. Archibald, 1930. For additional references see the notes by Professor Archibald p. 81 ff. in this second edition.

* Bradley: *American Mathematical Monthly*, Nov. 1925, p. 469.

ANSWERS

ANSWERS

§1.3

1. $13/11$
2. $15/8$
3. $13/23$
4. $25/12$
5. $12/19$
6. $24/13$
7. $11/27$
8. $43/48$
9. $x - 1$
10. $x^2 + 1$
11. $x - 1$
12. $(x + 1)^2$
13. $x^2 + 3x + 2$
14. $x^2 - 2x + 4$
15. $x^2 - 1$
16. $(x - 1)^2$
17. $x + 1$
18. $x^2 + 5x + 1$
19. $(x^2 + 3x + 1)^2$
20. $k = -5$
21. $k = \pm 1$
22. $27a^2d^2 + 4ac^3 - 18abcd + 4b^3d - b^2c^2$
23. $k = 0$, or 1

§1.5

1. $-1, -5$
2. $-2/3, -3/2$
3. $1, -5/2$
4. $-3, -3$
5. $\frac{1}{2}(1 \pm \sqrt{-3})$
6. $\frac{1}{8}(5 \pm \sqrt{-7})$

§2.2

1. 5
2. 16
3. 26
4. 10
5. 0
6. 0
7. 0
8. 0
9. $\cos^2 x$
10. 1
11. 1
12. 1
13. $(5, 2)$
14. $(4, 3)$
15. $(2, -3)$
16. $(3, -5)$
17. $(\frac{1}{3}, \frac{1}{2})$
18. $(\frac{10}{19}, \frac{5}{26})$
19. $(\frac{2}{3}, -\frac{3}{2})$
20. $(2, 3)$

§2.3

1. -57
2. 101
3. 0
4. 34
5. 80
6. 0
7. 0
8. 0
9. $(5, 2, 3)$
10. $(4, -2, 1)$
11. $(\frac{7}{13}, -\frac{6}{13}, \frac{12}{13})$
12. $(+1, +2, +3)$
13. $(3, -1, 2)$
14. Det. $= 0$
15. $(-1, 2, \frac{1}{2})$
16. $(10, 2, -6)$

§2.5

1. Case V.
2. Case I: Three planes meet in a common point $(3, 2, 1)$.
3. Case V: Two parallel planes cut by a third plane.
4. Case IV: Add the last two and divide by two.

5. Case II. **6.** Case III. **7.** Case IV.
8. Case V: The three planes intersect in three parallel lines.
9. Case IV. **10.** Case V (2).

§3.3

1. $8 + 8i$
2. $8 + 2i$
3. 8
4. $6i$
5. $25i$
6. $13i$
7. $-56 + 40i$
8. $(x^2 - y^2) + 2xyi$
9. $16 + 30i$
10. $16 - 30i$
11. $-9 + 46i$
12. $-46 + 9i$
13. $\frac{11}{5} + \frac{7}{5}i$
14. $\frac{1}{13}(-3 + 28i)$
15. i
16. $\frac{1}{13}(19 + 4i)$
17. $\frac{1}{29}(-1 + 41i)$
18. $\frac{1}{17}(23 + 7i)$
20. (a) $(5, 2)$; (b) $(7, 3)$; (c) $(2, 3)$; (d) $(3, 5)$, and their negatives.
21. (a) $3 - 2i$; (b) $4 + 3i$; (c) $5 + 4i$, and their negatives.

§3.4

2. (a) $\sqrt{13}$, arctan $\frac{3}{2}$;
(b) $\sqrt{13}$, arctan $\dfrac{-3}{2}$;
(c) 5, arctan $\frac{3}{4}$;
(d) 5, arctan $\dfrac{-4}{3}$;
(e) 10, arctan $\frac{4}{3}$;
(f) 13, arctan $\dfrac{-5}{12}$;
(g) 5, $90°$;
(h) 5, $270°$
(i) $\sqrt{13}$, arctan $\dfrac{2}{-3}$;
(j) 1, $60°$
(k) 1, $-60°$;
(l) 1, $135°$.

4. (a) $-\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$;
(b) $\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$;
(c) $3 \pm 2i$;
(d) $-3 \pm 2i$;
(e) $2 \pm 3i$;
(f) $-2 \pm 3i$;
(g) $1, -\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$;
(h) $\pm 1, \pm i$;
(i) $-1, \dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$;
(j) $\pm 1, -\dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i, \dfrac{1}{2} \pm \dfrac{\sqrt{3}}{2}i$.

§3.7

1. (a) $3(\cos 0° + i \sin 0°)$;
(b) $2(\cos \pi + i \sin \pi)$;
(c) $3(\cos 90° + i \sin 90°)$;
(d) $4(\cos 270° + i \sin 270°)$;
(e) $\sqrt{2}(\cos 45° + i \sin 45°)$;
(f) $\sqrt{2}(\cos 315° + i \sin 315°)$;
(g) $2(\cos 60° + i \sin 60°)$;
(h) $2(\cos 300° + i \sin 300°)$;
(i) $2(\cos 120° + i \sin 120°)$;
(j) $2(\cos 330° + i \sin 330°)$;
(k) $2(\cos 30° + i \sin 30°)$;
(l) $2(\cos 240° + i \sin 240°)$.

2. $\frac{1}{2}(\sqrt{3}+i)$
3. $\frac{1}{2}(1+\sqrt{3}\,i)$
4. -1
5. $-\frac{1}{2}+\frac{\sqrt{3}}{2}i$
6.
7. $\frac{\sqrt{2}}{2}(1+i)$
8. 1
9. 1
10. $\frac{1}{2}(1+\sqrt{3}\,i)$
11. $\frac{\sqrt{2}}{2}(1+i)$
12. $\frac{1}{2}(\sqrt{3}-i)$
13. -1
14. i
15. i

16. (a) 9;
(b) 16;
(c) $4(-1-i)$;
(d) $2(-1-i)$;
(e) $8(-1+\sqrt{3}\,i)$;
(f) $16(\sqrt{3}+i)$;
(g) -1728;
(h) -64;
(i) $128(-1-\sqrt{3}\,i)$;
(j) 1;
(k) 1;
(l) -1;
(m) $-9\sqrt{3}\,i$;
(n) $-117-44i$;
(o) $-239+28560i$.

§3.10

1. $\frac{\sqrt{2}}{2}(1+i), \frac{\sqrt{2}}{2}(-1-i)$
2. $\frac{\sqrt{2}}{2}(-1+i), \frac{\sqrt{2}}{2}(1-i)$
3. $\frac{\sqrt{2}}{2}(1\pm i), \frac{\sqrt{2}}{2}(-1\pm i)$
4. 1, $\cos\theta+i\sin\theta$, $\theta = 72°, 144°, 216°, 288°$
5. -1, $\cos\theta+i\sin\theta$, $\theta = 36°, 108°, 252°, 324°$.
6. ±1, $\frac{1}{2}(1\pm\sqrt{3}\,i)$, $\frac{1}{2}(-1\pm\sqrt{3}\,i)$
7. $-i$, $\frac{1}{2}(\pm\sqrt{3}+i)$
8. i, $\frac{1}{2}(\pm\sqrt{3}-i)$
9. $\cos 22\overset{\circ}{.}5 + i\sin 22\overset{\circ}{.}5$; $-\cos 22\overset{\circ}{.}5 - i\sin 22\overset{\circ}{.}5$; $\cos 112\overset{\circ}{.}5 + i\sin 112\overset{\circ}{.}5$; $-\cos 112\overset{\circ}{.}5 - i\sin 112\overset{\circ}{.}5$
10. $\cos 22\overset{\circ}{.}5 - i\sin 22\overset{\circ}{.}5$; $-\cos 22\overset{\circ}{.}5 + i\sin 22\overset{\circ}{.}5$; $\cos 67\overset{\circ}{.}5 + i\sin 67\overset{\circ}{.}5$; $-\cos 67\overset{\circ}{.}5 - i\sin 67\overset{\circ}{.}5$
11. ±1; $\pm i$; $\frac{\sqrt{2}}{2}(\pm1\pm i)$
12. $\pm i$; $\frac{1}{2}(\pm\sqrt{3}\pm i)$
13. $\cos\theta+i\sin\theta$, $\theta = 10°, 130°, 250°$
14. $\pm\frac{1}{2}(\sqrt{3}+i)$, $\pm\frac{1}{2}(-1+\sqrt{3}\,i)$
15. $\cos\theta+i\sin\theta$, $\theta = 30°, 102°, 174°, 246°, 318°$
16. $\cos\theta+i\sin\theta$, $\theta = 50°, 110°, 170°, 230°, 290°, 350°$
17. $\sqrt[4]{2}\,(\cos\theta+i\sin\theta)$, $\theta = 22°.5, 202°.5$
18. $\sqrt[4]{2}\,(\cos\theta+i\sin\theta)$, $\theta = 157°.5, -22°.5$
19. $\sqrt[10]{2}\,(\cos\theta+i\sin\theta)$, $\theta = 45°, 117°, 189°, 261°, 333°$

20. $\sqrt[3]{2}\,(\cos\theta + i\sin\theta)$, $\theta = 20°, 140°, 260°$
21. $\sqrt{2}\,(\cos\theta + i\sin\theta)$, $\theta = 75°, 255°$
22. $\pm\dfrac{\sqrt{2}}{2}(\sqrt{3} - i)$
23. $\pm\frac{1}{2}(\sqrt{3} - i)$
24. $\cos\theta + i\sin\theta$, $\theta = 105°, 225°, 345°$
25. $\cos\theta + i\sin\theta$, $\theta = 75°, 195°, 315°$
26. $3, \frac{3}{2}(-1 \pm i\sqrt{3})$
27. $\dfrac{3\sqrt{2}}{2}(\pm 1 \pm i)$
28. $2\,(\cos\theta + i\sin\theta)$, $\theta = 0°, 72°, 144°, 216°, 288°$
29. $\pm 2i, \pm\sqrt{3} \pm i$
30. $\pm 2, \pm 1 \pm i\sqrt{3}$
34. 0 and 3
35. $2, 0, -2$
36. $\omega, \omega^2, \omega^2, 1, \omega$

§4.2

1. 12	**11.** 2	**21.** 0.8	**31.** 0.001
2. 3.5	**12.** 2.8	**22.** 0.003	**32.** 0.01
3. 3.2	**13.** 3	**23.** $\frac{8}{13}$	**33.** $\frac{1}{6}$
4. 3	**14.** 3	**24.** 0.0002	**34.** 0.2
5. 4	**15.** 4	**25.** 0.004	**35.** 0.1
6. 501	**16.** 1.25	**26.** 0.0001	**36.** 0.02
7. 2	**17.** 0.5	**27.** 0.001	**37.** 0.03
8. 1.01	**18.** $\frac{4}{7}$	**28.** 0.0001	**38.** 0.04
9. 2	**19.** 0.2	**29.** 0.00001	**39.** 0.05
10. 2	**20.** 0.5	**30.** 0.0001	**40.** $\frac{1}{3}$

§4.3

1. $4x^3 + 9x^2 + 10x$
$12x^2 + 18x + 10$
2. $5x^4 - 21x^2 + 6x$
$20x^3 - 42x + 6$
3. $15x^4 + 20x^3 - 6x^2 + 4$
$60x^3 + 60x^2 - 12x$
4. $8x^3 + 9x^2 - 8x + 5$
$24x^2 + 18x - 8$
5. $3x^2 - 12x + 11$
$6x - 12$
6. $3x^2 - 2x$
$6x - 2$
7. $3x^2 - 14x + 12$
$6x - 14$
8. $4x^3 - 10x$
$12x^2 - 10$
9. $4x^3 + 9x^2 + 8x + 12$
$12x^2 + 18x + 8$
10. $5x^4 + 8x^3 - 18x^2 - 24x - 15$
$20x^3 + 24x^2 - 36x - 24$

§4.5

1. -3	**5.** -115	**9.** -1	**13.** -123
2. 12	**6.** 73	**10.** 6	**14.** 3
3. 1	**7.** 12	**11.** -4	**15.** -200
4. 32	**8.** 3	**12.** -6	**16.** 24

§4.6

1. 5	**3.** 4	**5.** 11	**7.** 7
2. 12	**4.** 5	**6.** 5	**8.** 10

§4.7

1. $x^3 - x^2 - 9x - 13, R = -23.$
2. $x^4 + 2x^3 - 9x^2 + 20x - 43, R = 91$
3. $x^4 + 3x^3 + 4x^2 + 15x + 45, R = 128.$
4. $2x^2 + 4x - 1, R = 4$
5. $2x^3 + x^2 - x, R = -7$
6. $2x^4 + x^3 - 6x^2 - 12x - 19, R = -48$
7. $3x^3 + x^2 + x + 2, R = 9$
8. $2x^3 + x^2 + x + 2, R = 1$
9. $x^4 - 2x^3 + 4x^2 - 8x + 16, R = 0$
10. $x^5 + 2x^4 + 4x^3 + 8x^2 + 16x + 32, R = 0$
11. $3x^3 - 2x^2 + x - 10, R = 33$
12. $x^4 + x^2 - x - 2, R = 1$

§4.9

1. Min. $y = -1$ at $x = 0$.
2. Min. $y = -2$ at $x = 0$.
3. Min. $y = \frac{7}{4}$ at $x = \frac{3}{2}$.
4. Min. $y = \frac{3}{4}$ at $x = \frac{3}{2}$.
5. Min. $y = -\frac{1}{4}$ at $x = \frac{3}{2}$.
6. Min. $y = -\frac{5}{4}$ at $x = \frac{3}{2}$.
7. Min. $y = -\frac{9}{4}$ at $x = \frac{3}{2}$.
8. Min. $y = -\frac{13}{4}$ at $x = \frac{3}{2}$
9. No max. or min.
10. Max. $y = 2$ at $x = -1$; min. $y = -2$ at $x = 1$.
11. Max. at $x = \dfrac{5 - \sqrt{13}}{3}$; min. at $x = \dfrac{5 + \sqrt{13}}{3}$.
12. Max. at $x = \dfrac{11 - \sqrt{19}}{3}$; min. at $x = \dfrac{11 + \sqrt{19}}{3}$.
13. Max. at $x = \dfrac{4 - \sqrt{19}}{3}$; min. at $x = \dfrac{4 + \sqrt{19}}{3}$.
14. Min. at $x = 0$, $x = \dfrac{15 + \sqrt{97}}{8}$; max. at $x = \dfrac{15 - \sqrt{97}}{8}$.
15. Min. $y = -27$ at $x = 3$.

16. Min. $y = 0$ at $x = 0$.
17. Min. $y = 0$ at $x = 1, 4$; max. $y = \frac{81}{16}$ at $x = \frac{5}{2}$.
18. Max. $y = 0$ at $x = 0$; min. $y = \dfrac{-256}{27}$ at $x = \dfrac{8}{3}$.
19. Max. $y = 32$ at $x = 0$; min. $y = 0$ at $x = 4$.
20. No max. or min.
21. Max. $y = 0$ at $x = 0$; min. $y = -256$ at $x = 4$.
22. Min. $y = 0$ at $x = 3$; max. $y = 26244/3125$ at $x = 9/5$.
23. Min. $y = 0$ at $x = 0$; max. at $x = \dfrac{9 - \sqrt{41}}{4}$; min. at $x = \dfrac{9 + \sqrt{41}}{4}$.
24. Max. $y = 8$ at $x = 0$; min. $y = -8$ at $x = \pm 2$.
25. Min. $y = 5$ at $x = 3$; max. $y = 9$ at $x = 1$.
26. Min. $y = -21$ at $x = 3$; max. $y = 104$ at $x = -2$.
27. Max. $y = -4$ at $x = 0$; min. $y = -20$ at $x = \pm 2$.
28. Max. $y = +1$ at $x = 0$; min. $y = -15$ at $x = \pm 2$.
29. Max. $y = 2$ at $x = 0$; min. $y = -14$ at $x = \pm 2$.
30. Max. $y = 16$ at $x = 0$; min. $y = 0$ at $x = \pm 2$.
31. Max. $y = -9$ at $x = 0$; min. $y = -25$ at $x = \pm 2$.
32. The roots are $3, \frac{1}{2}, \frac{1}{3}$. There is a max. between $\frac{1}{2}$ and $\frac{1}{3}$. Max. at $x = 0.41$; min. at $x = 2.14$
33. $-1, 2, 2$
34. $-2, -2, 3$
35. $2, 2, -3$
36. $3, 3, -2$
37. $5, 5, -2$
38. $-4, -4, 3$
39. $2, 2, 2, -4$

§5.2

1. $x^3 - 7x - 6 = 0$
2. $6x^3 - 17x^2 + 11x - 2 = 0$
3. $x^3 + 3x^2 - 4x - 12 = 0$
4. $x^4 - 8x^3 + 23x^2 - 28x + 12 = 0$
5. $x^4 - 9x^3 + 26x^2 - 24x = 0$
6. $x^4 - 3x^3 - 6x^2 + 28x - 24 = 0$
7. $24x^3 - 26x^2 + 9x - 1 = 0$
8. $6x^3 - 25x^2 + 32x - 12 = 0$
9. $24x^3 - 46x^2 + 29x - 6 = 0$
10. $9x^3 - 6x^2 - 53x - 30 = 0$
11. $x^3 - 3x^2 - 5x + 15 = 0$
12. $x^4 - 5x^2 + 6 = 0$
13. $x^3 - 11x^2 + 37x - 35 = 0$
14. $x^3 - 3x^2 - 3x + 1 = 0$

15. $x^4 - 10x^3 + 32x^2 - 34x + 7 = 0$
16. $x^3 - 3x^2 + 4x - 12 = 0$
17. $x^3 - 5x^2 + 9x - 45 = 0$
18. $x^3 - 10x^2 + 37x - 52 = 0$
19. $x^4 - 10x^3 + 50x^2 - 130x + 169 = 0.$
20. $x^4 - 8x^3 + 24x^2 - 44x + 35 = 0.$

§5.6

1. $x^3 - 9x^2 + 26x - 24 = 0$
2. $x^3 - x^2 - 8x + 12 = 0$
3. $x^4 - 10x^3 + 35x^2 - 50x + 24 = 0$
4. $x^3 - 2x^2 - 9x + 18 = 0$
5. $x^4 - 13x^2 + 36 = 0$
6. $x^4 - 7x^3 + 18x^2 - 20x + 8 = 0$
7. $x^3 - x^2 - 3x + 3 = 0$
8. $x^4 - 5x^2 + 6 = 0$
9. $x^5 - 15x^4 + 85x^3 - 225x^2 + 274x - 120 = 0$
10. $x^4 - 19x^2 - 6x + 72 = 0$
11. $x^4 - 4x^3 - 19x^2 + 46x + 120 = 0$
12. $x^4 - 4x^3 - 7x^2 + 22x + 24 = 0$
13. $x^4 - 4x^3 + x^2 + 8x - 6 = 0$
14. $x^4 - 6x^3 + 5x^2 + 18x - 24 = 0$
15. $x^4 - 7x^2 + 10 = 0$
16. $x^3 - 5x^2 + 11x - 15 = 0$
17. $x^4 - 13x^2 + 36 = 0$
18. $x^4 - 2x^3 - 13x^2 + 14x + 24 = 0$

§5.8

1. $\pm 1, 2 \pm \sqrt{3}$
2. $1, 2, 1 \pm i\sqrt{2}$
3. $-1, \pm i$
4. $1, 2, 3$
5. $2, 4, -1 \pm i\sqrt{3}$
6. $1, 3, 5, 7$
7. $2, 3, 2 \pm \sqrt{7}$
8. $1, 4, -2, -5$
9. $1, 5, -3, -7$
10. $\pm 1, 3, 5$
11. $1 \pm \sqrt{2}, 2, -1, -3$
12. $1 \pm i\sqrt{2}, 5, -1, -6$
13. $4, 2, -6$
14. $-3, -4, 5$
15. $1, 2, 3, -6$
16. $1, 2, 3, -2, -4$
17. $\pm 2, -3, -4, 7$
18. $\pm 1, \pm 2, \pm 3$
19. $\pm 1, \pm 2, \pm 5$
20. $\pm 1, \pm 2, \pm 3, \pm 5$
21. $3, \frac{1}{2}, -\frac{2}{3}$
22. No rational root
23. $5, 8, \frac{1}{3}$
24. $6, 7, -\frac{2}{3}$
25. $\frac{3}{2}, \frac{2}{3}, 5$
26. $5, 6, \frac{2}{3}$

§5.9

1. $-3, 2$
2. $2, 2, 3, -5$
3. $4, 5, 6$
4. $-3, 5, 8$
5. $-4, 2, 5, 8$
6. $-3, -10, 5, 8$
7. $\pm 2, \pm 9$
8. ± 8
9. $6, 8$
10. $2, 3$
11. ± 12
12. ± 6
13. ± 10
14. $\pm 3, \pm 8$
15. $-2, -5, 3, 4$
16. $-4, 2, 5, 6$
17. $3, 6, -4$
18. $\pm 2, \pm 3, -5$
19. $\pm 1, \pm 6, 5$
20. $\pm 6, 4$

§5.10

1. $-2, 1, 30$
2. $\pm 2, 5$
3. $5, 8$
4. $5, -6$
5. $1, 5, 8$
6. $1, 7, 10$
7. $-2, -4, 6, 11$
8. $\pm 5, \pm 6$
9. $\pm 2, \pm 5.$
10. $\pm 3, \pm 5, -3, -4.$

§5.11

1. $\frac{1}{2}, \frac{1}{3}, -2$
2. $\pm 1, \pm \frac{3}{2}$
3. $2, -\frac{1}{3}, -\frac{3}{2}$
4. $2, 3, -\frac{3}{4}$
5. $1, -3, -\frac{1}{3}$
6. $\pm \frac{1}{2}, 2, 3$
7. $2, 3, -1, \frac{1}{2}$
8. $\frac{3}{2}, \frac{2}{3}, -\frac{5}{2}$
9. ± 2
10. $\pm \frac{1}{2}, \frac{1}{3}$
11. $\pm \frac{1}{2}, \pm \frac{1}{3}, 1.$

§5.12

1. 1
2. -1
3. 1
4. -2
5. -1
6. -1
7. -2
8. -2
9. 2
10. $3, -2$
11. -3
12. 3
13. ± 1
14. $\pm i$
15. $1 \pm \sqrt{2}$
16. $\pm \sqrt{3}$
18. $G^2 + 4H^3 = 0$

§5.13

1. Double root at $x = 3, \frac{3}{2}$
2. One root less than 2; one greater than 5; one between 2 and 5.

In the intervals

3. $(0, 2), (2, 6), (6, +\infty).$
4. $(1, 2), (0, 1), (2, +\infty), (-1, -2).$
5. $(-\infty, -2), (0, 1), (1, 2), (2, +\infty).$
6. $(-\infty, -3), (-3, -1), (-1, 0), (1, 3), (3, +\infty).$
7. $(-\infty, -2), (-2, -1), (-1, 0).$
8. $(-\infty, -6), (-6, 2), (2, 4), (4, \infty).$

§6.2

1. $x^3 + 5x^2 - 4x - 3 = 0$
2. $x^3 + 2x + 7 = 0$
3. $x^3 - 1 = 0$
4. $4x^3 + 2x^2 + 3x - 5 = 0$
5. $2x^4 + 3x^2 - 7 = 0$
6. $3x^4 + 4x^3 - 2x - 5 = 0$
7. $\quad + x^4 + 2x^3 + 3x^2 + 5x + 4 = 0$
8. $5x^6 + 3x^3 + x^2 - 7 = 0$
9. $7x^4 + 4x^3 + 5x^2 - 3x - 2 = 0$
10. $x^5 + x^2 + x + 3 = 0$
11. $3x^6 + 5x^3 - 7x - 5 = 0$
12. $2x^5 + 3x^4 - 5x^2 + 6 = 0$
13. $5x^3 - 2x^2 + 12x - 8 = 0$
14. $x^4 - 3x^3 + 20x - 48 = 0$
15. $x^4 - x^3 + 3x^2 - 18x + 27 = 0$
16. $x^4 - 8x^2 + 80 = 0.$
17. $x^3 + 4x^2 + 16x + 64 = 0$
18. $3x^4 - 5x^3 - 7x - 2 = 0$
19. $x^3 - 6x^2 - 8x + 32 = 0$
20. $x^4 + 15x^2 + 25x + 250 = 0$
21. $x^6 + 32x + 64 = 0$
22. $x^5 + 45x^2 - 81 = 0$
23. $x^3 + 3x^2 + 10x - 100 = 0$ (5)
24. $x^3 + 3x^2 - 12x + 54 = 0$ (6)
25. $x^4 - 5x^3 + 18x - 27 = 0$ (3)
26. $x^3 + 2x^2 - 9x + 48 = 0$ (6)
27. $x^4 - 20x^2 + 100x - 1250 = 0$ (10)
28. $x^4 - x^3 + 2x^2 - 35 = 0$ (5)
29. $x^4 + 3x^3 + 8x^2 + 16x + 64 = 0$ (4)
30. $x^4 - 10x^2 + 12x - 32 = 0$ (2)
31. $x^5 - 3x^4 + 12x^3 + 54x^2 - 1152 = 0$ (6)
32. $x^6 - 25x^3 + 250x - 375 = 0$ (5)
33. $x^4 - 5x^3 - 15x^2 + 135x + 360 = 0$ (30)

§6.3

1. $x^2 + 2x + 1 = 0$
2. $x^3 + 1 = 0$
3. $4x^3 - 7x^2 + 5x + 3 = 0$
4. $7x^4 - 5x^3 - 3x^2 + 2 = 0$
5. $2x^4 - 5x^3 - 5x + 2 = 0$
6. $2x^5 - 3x^4 + 5x^3 + 5x^2 - 3x + 2 = 0$
7. $5x^4 - 7x^3 + 7x - 5 = 0$

8. $x^5 + x^4 + 3x^3 + x + 1 = 0$
9. $5x^3 - 7x^2 + 3x + 1 = 0$
10. $3x^4 + 5x^3 - 5x - 3 = 0$

§6.4

1. $y^4 + 11y^3 + 43y^2 + 55y - 9 = 0$
2. $y^5 + 15y^4 + 94y^3 + 305y^2 + 507y + 353 = 0$
3. $4y^5 - 40y^4 + 158y^3 - 308y^2 + 303y - 129 = 0$
4. $3y^4 - 77y^3 + 720y^2 - 2876y + 4058 = 0$
5. $2y^4 + 11y^3 + 25y^2 + 20y = 0$
6. $y^4 + 2y^3 - 3y^2 = 0$
7. $y^4 - 10y^2 - 11y + 4 = 0$
8. $y^4 - 29y^2 - 59y + 5 = 0$
9. $y^3 - 27y - 2 = 0$
10. $y^3 - 12y - 3 = 0$
11. $y^4 + 3y^3 - 7y + 2 = 0$.

§6.5

1. $y^3 - 47y + 24 = 0$
2. $y^4 - 6y^2 + 1 = 0$
3. $y^3 - 12y - 9 = 0$
4. $y^4 - 24y^2 + 10 = 0$
5. $y^3 - 1 = 0$
6. $y^4 - 12 = 0$
7. $y^4 + y + 1 = 0$
8. $y^4 - 16y^2 - 40y - 29 = 0$
9. $y^5 - 6y^3 - 5y^2 + 2 = 0$
10. $y^5 + 30y^2 + 46y = 0$

§6.7

1. $1, 1, \frac{1}{2}(-3 \pm \sqrt{5})$
2. $1, \frac{1}{4}\{-1 \pm \sqrt{5} \pm i\sqrt{10 \pm 2\sqrt{5}}\}; \pm 1; \frac{1}{2}(\pm 1 \pm i\sqrt{3})$.
3. $\pm 1, 2, \frac{1}{2}, \frac{1}{2}(-3 \pm \sqrt{5})$
4. $1, 1, -1, -1, \omega, \omega^2$
5. $\frac{3}{2}, \frac{2}{3}, 2, \frac{1}{2}$
6. $2, 3, \frac{1}{2}, \frac{1}{3}$
7. $\pm 2, \pm 2, \pm\frac{1}{2}, \pm\frac{1}{2}$
8. $\pm 3, \pm 3, \pm\frac{1}{3}, \pm\frac{1}{3}$
9. $\pm 1, \omega, \omega, \omega^2, \omega^2, \frac{1}{2}(1 \pm i\sqrt{3}); \frac{1}{2}(1 \pm i\sqrt{3})$
10. $\frac{1}{2}(1 \pm i\sqrt{3}); \frac{1}{2}(-5 \pm \sqrt{21})$

11. $-1, 2 \pm \sqrt{3}, \frac{1}{2}(1 \pm i\sqrt{3})$
12. $1, 1, 1, -1, \frac{1}{4}(1 \pm i\sqrt{15})$
13. $2, \frac{1}{2}, -3, -\frac{1}{3}$.
14. $\pm 1, 2 \pm \sqrt{3}, 2 \pm \sqrt{3}, -2 \pm \sqrt{3}, -2 \pm \sqrt{3}$
15. $\pm 1, -1 \pm \sqrt{2}, -1 \pm \sqrt{2}, 1 \pm \sqrt{2}, 1 \pm \sqrt{2}$
16. $a = 1$: $0, -1, \frac{1}{2}(1 \pm i\sqrt{3})$
$a = 5$: $-1, \frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{8}(-3 \pm \sqrt{-55})$
$a = 16$: $-1, 1, 1, \frac{1}{3}(-1 \pm 2\sqrt{-2})$
$a = 45$: $-1, \frac{1}{8}(7 \pm i\sqrt{15}), \frac{1}{22}(-7 \pm i\sqrt{435})$
$a = \frac{1}{31}$: $-1, -2, -\frac{1}{2}, \frac{1}{6}(-5 \pm i\sqrt{11})$
17. $(x^2 + 1)(x^2 - \sqrt{3}x + 1)(x^2 + \sqrt{3}x + 1)$
18. $$\frac{-2 + \sqrt{2a^2 + 2a} \pm \sqrt{10a - 2a^2 - 4\sqrt{2a^2 + 2a}}}{2(1 - a)}$$
$$\frac{-2 - \sqrt{2a^2 + 2a} \pm \sqrt{10a - 2a^2 + 4\sqrt{2a^2 + 2a}}}{2(1 - a)}$$

§7.1

1. $-2, 1, 1$
2. $4, -2 \pm i\sqrt{3}$
3. $5, -\frac{5}{2} \pm i\frac{\sqrt{3}}{2}$
4. $4, -2, -2$
5. $3, 3, -6$
6. $9, \frac{3}{2} \pm i\frac{\sqrt{3}}{2}$
7. $2, 2, 5$
8. $3, -\frac{3}{2} \pm i\frac{\sqrt{15}}{2}$
9. $2, -1 \pm i\sqrt{6}$
10. $3, -\frac{3}{2} \pm i\frac{\sqrt{39}}{2}$
11. $5, -\frac{5}{2} \pm i\frac{\sqrt{147}}{2}$
12. $2, -1 \pm 3i$
13. $3, -3 \pm i\sqrt{6}$
14. $4, 1 \pm i$
15. $3, 3 \pm i$
16. $2, 2 \pm i$
17. $6, \pm i\sqrt{3}$
18. $4, -2 \pm i3\sqrt{3}$
19. $8, -1 \pm i4\sqrt{3}$
20. $4, \frac{11}{2} \pm i\frac{5\sqrt{3}}{2}$

§7.3

1. -351
2. 621
3. 81
4. -135
5. 0
6. 0
7. 0
8. 0
9. -175
10 -23
11. (1) 0
(2) $-18{,}252$
(3) -9747
(4) 0
(5) 0
(6) -9747
(7) 0
(8) -8640
(9) -5400
(10) $-35{,}100$
(11) $-1{,}271{,}403$
(12) $-11{,}664$
(13) $-42{,}336$
(14) -400
(15) -4
(16) -4
(17) $-18{,}252$
(18) $-428{,}652$
(19) $-3{,}195{,}072$
(20) $-33{,}075$

§7.4

1. 0.7422; −1.1371; 0.3949
2. $-2, 1 \pm \sqrt{3}$
3. $2 \cos 20°, 2 \cos 140°, 2 \cos 260°$
4. $3, 5, -2$
5. 3.8232; −2.9304; −0.8928
6. −1.4679; −4.8794; −2.6527
7. 0.9085; − 1.3747; − 0.5337
8. −5.0000; 0.4495; −4.4495

§7.8

1. $-1, -1, -1, -3; u = 4$
2. $-2, 4, -1 \pm i; u = -6$
3. $\frac{1}{2}(-1 \pm \sqrt{5}), \frac{1}{2}(1 \pm i\sqrt{11}); u = 2$
4. $2 \pm \sqrt{6}, 1 \pm i\sqrt{5}; u = 4$
5. $+2, -1, -1 \pm i\sqrt{2}; u = 1$
6. $-1, 3, -2 \pm \sqrt{5}; u = -4$
7. $1 \pm i, -3 \pm \sqrt{5}; u = 6$
8. $\pm 1, -2, -4; u = 2$
9. $\frac{1}{2}(-1 \pm i\sqrt{7}), \frac{1}{2}(-5 \pm \sqrt{41}); u = -2$
10. $2, 3, \frac{1}{2}(-5 \pm \sqrt{17}); u = 8$
11. $2, 4, -3 \pm \sqrt{19}; u = -2$
12. $1, 3, -2 \pm i; u = 8$
13. $2, 4, -3 \pm \sqrt{13}; u = 4$
14. $2, 2, -2 \pm \sqrt{12}; u = -4$
15. $-2, 6, -2 \pm i\sqrt{2}; u = -6$
16. $2 \pm \sqrt{10}, -2 \pm i2\sqrt{2}; u = 6$
17. $2, 4, -3 \pm \sqrt{11}; u = 6$
18. $-2, -2, 2 \pm i2\sqrt{2}; u = 16$
19. $\pm 3, -3 \pm 2i; u = 4$
20. $\pm 4, -3 \pm i\sqrt{5}; u = -2$
21. $1, -1, -1, -1; u = 0$
22. $1, 2, -2, -3; u = 8$

§8.1

	Th. 3 gives	The largest root is		Th. 3 gives	The largest root is
1.	5.69	4	16.	4.5	3
2.	9.72	5	17.	3	1
3.	9	8	18.	5.2	3

4.	13	12	**19.**	5.3	3
5.	11.3	10	**20.**	2.36	1
6.	9.55	8	**21.**	3.55	1
7.	6.2	4	**22.**	3.52	2
8.	33	30	**23.**	3	2
9.	8.8	5	**24.**	4	3
10.	2.42	1	**25.**	5.217	3
11.	11.2	10	**26.**	5	4
12.	11.2	10	**27.**	5	4
13.	6	3	**28.**	4	3
14.	5.47	4	**29.**	4	2.1+
15.	3.65	2	**30.**	4	3

Exercise	1	2	3	4	5	6	7	8	9	10
Th. 3	5	6	11	111	5	5	3	3	3	81
Th. 4	12	6	2	6	12	17	6	3	6	5

§8.5

1.

a	b	c	d	e	f	g	h	i	j
3	1	4	3	5	4	5	5	4	5

§8.6

1. 5 **2.** 2 **3.** 3 **4.** 3

§10.5

1. 1.15 **5.** 1.80194 **9.** 0.739 **13.** −3.183
2. 4.06 **6.** 0.28466 **10.** 0.947 **14.** 2.25
3. 1.347 **7.** 4.2644 **11.** 1.895 **15.** 0.48
4. 2.25132 **8.** 0.31469 **12.** 1.30
16. −0.67; 1.42; 5.25 **17.** 1.35; 4.06
18. −2.67; −0.58; 3.25

§11.6

1. $+, +, +,$ **2.** $-, -, +, +$

3. $a_{14}a_{21}a_{33}a_{42} + a_{14}a_{22}a_{31}a_{43} + a_{14}a_{23}a_{32}a_{41} - a_{14}a_{21}a_{32}a_{43} - a_{14}a_{22}a_{33}a_{41} - a_{14}a_{23}a_{31}a_{42}$

4. $a_2b_1c_4d_3e_5 - a_2b_1c_4d_5e_3$

5. $a_1b_2c_3(d_4e_5f_6 - d_4e_6f_5 + d_5e_6f_4 - d_5e_4f_6 + d_6e_4f_5 - d_6e_5f_4)$

§11.17

1. (5, 3, 2)
2. (5, 4, 1)
3. (6, 2, 4)
4. (3, 1, 2, −1)
5. (5, 3, 1, −1)
6. (2, −1, 1, 2)
7. (3, −2, 1, 2)
8. (2, −2, 3, 1)
9. (3, −1, 2, −1)

§13.5

1.

a	b	c	d	e	f	g	h
−16	−147	−31	−256	256	0	68	30,685

2. The discriminant in every case is zero. The roots are:

(a) 1, 1, −4
(b) −4, −4, 1
(c) 1, 1, −3
(d) 2, 2, $\frac{1}{2}(-1 \pm i\sqrt{3})$
(e) 1, 1, $\frac{1}{2}(-1 \pm i\sqrt{3})$
(f) 2, 2, −3, −1
(g) −3, −3, 2, 6
(h) 2, 2, 0, −4

§14.8

1. $x^4 - 4x^2 + 1 = 0$
2. $x^4 - 10x^2 + 1 = 0$
3. $x^2 - 6 = 0$
4. $x^4 - 10x^2 + 22 = 0$
5. $x^4 - 16x^2 + 4 = 0$
6. $x^4 - 8x^2 + 20 = 0$

§14.11

1. $x^4 - 10x^2 + 18 = 0$
2. $x^4 - 14x^2 + 9 = 0$
3. $x^4 - 12x^2 + 20 = 0$
4. $x^8 - 40x^6 + 352x^4 - 960x^2 + 576 = 0$
5. $x^8 - 12x^6 + 6x^4 - 12x^2 + 1 = 0$
6. $x^8 - 8x^6 - 180x^4 - 16x^2 + 4 = 0$
7. $x^8 - 12x^6 + 48x^4 - 72x^2 + 34 = 0$
8. $x^8 - 32x^6 + 260x^4 - 224x^2 + 4 = 0$
9. $x^{16} + 16x^{14} + 100x^{12} + 304x^{10} + 452x^8 + 288x^6 + 64x^4 - 2 = 0$
10. $x^8 + 16x^7 + 92x^6 + 208x^5 + 16x^4 - 640x^3 - 760x^2 + 32x + 292 = 0$
11. $x^8 + 8x^7 + 12x^6 - 40x^5 - 102x^4 + 8x^3 + 68x^2 - 72x - 73 = 0$
12. $x^4 - 4x^3 - 4x^2 + 16x - 8 = 0$
13. $x^4 - 4x^3 - 8x^2 + 24x - 4 = 0$

INDEX